二十世纪西方哲学译丛

理性、真理与历史

Reason, Truth and History

[美]希拉里·普特南著
童世骏 李光程译

上海译文出版社



二十世纪西方哲学译丛





理性、真理与历史

Reason, Truth and History

〔美〕希拉里·普特南著

上海译文出版社

Hilary Putnam

REASON, TRUTH AND HISTORY

Published by the Press Syndicate of the University of Cambridge

本书根据英国剑桥大学出版社 1981 年版译出

理性、真理与历史

〔美〕希拉里·普特南 著

童世骏　李光程 译

上海译文出版社出版、发行

上海延安中路955弄14号

全国新华书店经销

上海市印刷十二厂印刷

开本 850×1168　1/32　印张 8　插页 3　字数 188,000

1997 年 2 月第 1 版　1997 年 2 月第 1 次印刷

印数：0,001－5,000 册

ISBN 7-5327-1710-0/B·083

定价：15.00 元

《二十世纪西方哲学译丛》选收本世纪西方哲学界各主要流派影响较大的著作。通过有选择的译介，旨在增进文化积累，拓展学术视野，丰富研究课题，为了解和研讨现代西方哲学提供系统而完整的第一手资料，以利于我国理论界、学术界深化对西方文化的借鉴和批判。

译 者 的 话

《理性，真理与历史》的作者希拉里·普特南是当代美国“后分析哲学”或“后实证主义哲学”的主要代表之一，哈佛大学的哲学教授。普特南1926年出生于一个犹太作家家庭。这种家庭背景对他哲学的影响有两个方面。一方面，普特南得自其父的文学兴趣使他即使在讨论高度技术化的哲学问题时，也能深入浅出，文笔晓畅。另一方面，普特南——尤其在其后期——在哲学著述中表现了强烈的神学兴趣，设法对希伯来传统作新的理解。在他看来，宗教的意义在于使人类意识到自己的局限性和可错性，同时又表达了人类追求的解放和得救的理想。如果说这种理想在普特南后期的思想中具有较为直接的宗教色彩的话，在其年轻时期，它则表现为激进的左翼立场。普特南积极参加60年代美国反战运动和学生造反运动，曾担任著名的“学生争取民主社会”组织的哈佛教师代表和一个叫做“进步劳动党”的毛派组织的成员。普特南的这些活动同他对于哲学的本性的看法有密切联系。在他看来，哲学虽不能拯救世界，也不能仅限于解答逻辑或语言疑难。在大学时，普特南就对基尔凯郭尔、弗洛伊德、尤其是马克思极感兴趣。但是，在分析哲学一统天下的当时的美国哲学界，他通向哲学教授的唯一途径是接受严格的分析哲学训练。这种训练之“严格”有两重意思。一方面，他必须强迫自己把“分析哲学”当成“哲学”的同义词。另一方面，他必须精通此派哲学的思路、方法和术语。他后来对分析哲学“揭竿而起”，表明这种“严格训练”在前一方面没有成功，但他后来在语

言哲学、逻辑哲学、数学哲学、精神哲学、认识论、伦理学等各个领域的卓越成就，则表明他从自己受的分析哲学训练中获益匪浅，因为只有在哲学上充分地"知己知彼"，才能在这些长期由分析哲学独霸一方的领域攻城夺寨。

在普特南多产的哲学生涯中，《理性，真理与历史》是其代表作之一。它自 1981 年出版以来经过短短几年时间便有了当代经典的声望。在英、美哲学界的影响自不待言，这里只提两件事情。其一，1982 年 12 月，美国哲学学会（APA）曾为此书召开专题讨论会，这并不是常见的事。其二，笔者之一曾有意翻阅此书在剑桥大学图书馆和剑桥哲学系的出借记录，结果是几乎从未有过间断。尤为意味深长的是，在一些欧洲大陆国家，此书同样受到青睐。笔者着手翻译此书就是受到一位留学德语国家的学者的启发，据他说这本书在他们那里是哲学系学生的十本必读书之一。这近乎耸人听闻。因为按知名度排列，古今大师何止上百，何以普特南能跻身其间？但无论如何，这说法还是诱惑着我们费时近一年完成了此书的中译本。

《理性，真理与历史》一书的内容涉及语言哲学、逻辑哲学、精神哲学、科学哲学、历史哲学、形而上学、知识论和伦理学等哲学分支的大量哲学问题，尤其对指称和意义、真理的本质、合理性和历史等问题提出了独到的见解。在序言中，普特南谈到本书的主旨在于打破一系列长期支配哲学家及普通人的二分法思维方式。其中最主要的是关于真理和理性的主客观的二分法，也就是关于真理的符合论和彻底相对主义之间的二分法。他试图说明，许多意识形态并非非此即彼。因此我们可以拒绝符合论而无须成为相对主义者，或者，我们可以摒弃相对主义而无须接受真理的符合论。

首先我们来看看所谓的外在的实在论和内在的实在论。

按普特南的说法，符合论真理观实际上是一种形而上学实在论，或外在的实在论。其基本观点有三：(1)"世界由独立于心灵的

客体的固定总体组成";(2)"对世界存在的方式恰好只有一个真实而全面的描述";(3)"真理包含语词之间或思想记号和外部事物之间的某种符合(对应)关系"。

很显然,这种观点肯定世界是外在于心灵的,因而是外在论或客观论;它肯定真理只有一个,且真理只是一个陈述符合(或对应)一个外部事实,因而是符合论或唯一真理论;又由于这种理论离开人所制定的理论来谈论实在,因而是形而上学实在论,或用一个形象的说法,这是从一种超然于人的"上帝的眼光"来看世界和实在。

相对主义者走另一个极端。他们把一切思想和观点都看作是主观的。真理或正当性只是相对于个人或团体而言的。用法伊尔阿本德著名的话说便是:"怎么都行",没有绝对的、不变的标准。

在普特南看来,这两种倾向都是不可接受的。如果外在实在论是正确的,我们便可以设想一个离开我们语言和理论的实在,这也就意味着我们可以设想如下情形:我们实际上都只是一些与世隔绝的营养钵里的大脑,因而我们所谈论的实际上都只是一些幻觉。但是这个假说是自相矛盾的:一旦我们能够考虑它是真的还是假的,我们的思想就必定能够真正地指称外部对象,这就表明了这个假说是不成立的。因为,任何能想到他自己是一个钵中之脑的人,他就并不是一个钵中之脑。所以,尽管在某一可能世界之上的人能想和说我们所想和说的话,但他决不能指称我们可指称的东西,尤其是,他们不能说他们是钵中之脑。

相对主义也是站不住脚的。普特南的推论是,在坚持一个观点的同时又宣称没有其他观点是同样值得坚持的,这似乎在逻辑上自相反对。普特南引用了维特根斯坦对相对主义的诘难。这个诘难是,相对主义者不能区别是对的和认为是对的。如果看上去是对的就是对的,则无所谓对错了,甚至连自己的标准的对错也没有了。结果一个彻底的相对主义者不能肯定任何事情,也不能投身任何前后一致的实践,所以成为一个相对主义者等于"某种精神自

系”。

剩下来的唯一选择在普特南看来只能是内在的实在论。

根椐普特南的解释，所谓内在论是说：(1)只有从一个理论或描述之内问“世界是由什么对象组成”这个问题才是有意义的；(2)“对世界的‘真实’的理论或描述不止一个”；(3)“‘真理’是某种(理想化的)合理的可接受性”(引自本书第56页)。

这是一种与上帝的眼光相对的理论。它认为我们不能离开某一理论来谈论实在，也就是说，我们所谈论的实在并不外在于我们在其中谈论的理论框架。对此，普特南利用现代科学的事实来论证。现代科学表明，不同理论中的似乎不相容的术语常常可以对应于同一实在对象(当然不是在同一个理论内对应)。例如，这个理论中的被称为“点”的对象在另一个理论中可被称为“收敛的线段集”。同时，同样的术语也可以在不同的理论中对应不同的对象。这个事实表明，世界本身具有容许不同描述、对应、“绘制”的性质，而我们不能假定有一种“上帝的眼光”，离开任何人所制定的理论来谈论实在本身(物自体)如何如何。

在普特南看来，坚持内在论，我们既能相信真理，相信有真理的标准，从而避免相对主义，又能通过坚持有多种世界，有多种世界的存在方式，有多种理论和真理，其中有正确的，也有错误的，没有一成不变的标准，从而避免形而上学实在论或外在实在论的单一世界和唯一真理。

现在让我们转向指称理论。

很显然，对形而上学实在论的“上帝的眼光”和“钵中之脑”的假设的拒绝意味着对一种神秘的指称理论的排斥。所谓神秘的指称理论，是指这样的一种理论，它认为名称和它的承担者，概念和指称物之间有一种必然的、神秘的联系。这种观点从洛克开始就已形成，其理论基础在于，(1)意义在于头脑之中，知道一个概念等于拥有一种想象；(2)内涵或意义决定真值条件，语词内在地表征或

指称它们所表征或指称的东西。普特南指出，一旦人们意识到名称和其承担者只有一种上下文的、偶然的、约定俗成的联系，就很难明白为什么名称的知识应该有任何神秘的意义。为此他和克里普克一样，认为名称所以能指称对象，是因为有一条因果联系的链条将名称与指称联系在一起。因此，语词所指称的不是由心灵或意向所决定，而是由一个语言共同体所建立的范式和存在于超语言的世界中的历史（因果联系）所决定。指称一个对象包含大量非语言因素，特别是我们与我们所能谈论的事物之间的因果联系。这样，我们就能说明为什么有时各人使用的概念是不同的，但通过一定的历史的、社会的链条，常常把不同的概念引导到同一指称上来。这在科学理论中尤为明显。例如，尽管许多科学研究者对电子的描述与玻尔的描述不同，但这个名称指的对象是相同的。这是因为有一条历史的、社会的传递链条，把不同研究者与这个名词所指称的对象连在一起。如果依神秘的指称理论，这种不同术语指称同一对象的情况便不会发生，从而科学知识的会聚这一普遍现象也变得不可理解。

指称理论运用到科学哲学中的结果，使得普特南认为，成熟科学理论中的名词具有指称，这也等于承认分子、电子等理论实体的存在，这也是普特南有时被人称为科学实在论的原因（实际上他本人并不承认，而宁可称自己为非实在论者）。

我们再来看普特南的另一个主题——真理观。前面已提到，从内在论出发，普特南强烈反对符合论的真理观。普特南的意思是，理论中谈论的对象都具有洛克所谓的“第二性质”，因而都具有主观成分。不是先有一个独立的外在的对象，然后有一个对之加以摹拟的理论即真理。因为有时不相容的理论实际上都可以相互翻译，而按照真理与外部实在一一对应的观点，这是不容许的。而且，普特南继续诘问，如何比较理论与理论之外的实在以确定两者是否符合呢？塔斯基的真理论曾被认为是为实在论或真理的符合论提

供了依据。普特南否认这种看法。在他看来，塔斯基的真理定义在哲学上是中立的，只是一种"去掉引号的真理观"。因为根据任何真理论，"雪是白的"都等值于"'雪是白的'是真的"。那么普特南自己的真理观又是怎样的呢？普特南提出了一个重要思想，即应把真理问题与合理性问题结合起来解决。"在真理概念和合理性概念之间有着极其密切的联系。粗略说来，用以判断什么是事实的唯一标准就是什么能合理地加以接受"（序言，第 2 页）。"'真理'是某种（理想化的）合理的可接受性"（见本书第 56 页）。

这里先要对合理性问题的背景稍加交待。合理性问题可以说是当今科学哲学的中心课题。正如石里克、瓦托夫斯基等人所说，科学哲学的使命在于理解科学。在 19 世纪实证主义者穆勒、马赫等人那里，所谓理解科学就是理解科学家进行研究时的心理活动及其要素。随着自弗雷格以来的反心理主义潮流逐渐占据优势，科学哲学的主流——逻辑经验主义——的哲学家们实际上把知识的逻辑结构，逻辑方法和规则等问题作为"理解科学"的主要角度。波普尔站在逻辑主义的立场上（认为知识论是"没有主体的"）提出了尔后历史主义学派的科学哲学的中心课题：科学知识的历史发展。历史主义学派的科学哲学的演化结果是出现了一个与逻辑主义尖锐对立的极端：逻辑主义强调科学的合理性并把它等同于精确的逻辑方法、规则或程序，历史主义科学哲学家如库恩、法伊尔阿本德则否定科学中存在着前后同一的、普遍接受的逻辑方法、规则或程序，并以此为借口否认了科学的合理性本身。这样，一门以"理解科学"为己任的哲学部门居然得出了"科学没有合理性"即"科学不可理解"的结论。难怪法伊尔阿本德要自诩是科学哲学的掘墓人。正是在这种情况下，"科学到底有没有合理性？如果有的话，表现在什么地方？"这两个问题，成了进入 60 年代尤其是 70 年代以来科学哲学的中心问题。在本书中，普特南就几次提到真正的哲学家的任务就是研究合理性问题。在回答这个问题时，一群被称作"新历

史主义学派"的科学哲学家力图既承认科学的合理性，又承认科学的历史性。其主要办法是拆除横在科学理论与科学方法之间的樊篱，认为科学的合理性之所系的方法并不外在于理论，而是从理论中得来，并随着理论的变化而变化。普特南就是这群哲学家的主要代表之一。

但是，普特南并不认为合理性问题仅限于科学，尤其是精密科学领域。事实上，他再三强调，应当扩大对于"合理性"的理解，把道德、政治、文学和历史等等领域也包括在内。从这样的观点出发，普特南把合理性问题分成两个小问题，一是"善是不是合理的？"一是"合理的为何是善的？"在搞清"合理的"是什么意思之前，前一个问题是得不出正确答案的。如现代一些学者认为"合理的"只可针对手段而言，不能针对目的而言，他们自然认为目的的善与恶同合理与不合理无关。因此，重要的是回答第二个问题。普特南认为，被认为是合理性之标志的"自洽"、"简单"、"可检验"、"有预见力"等等，和道德领域中与"善"有关的"正当"、"恰当"等等一样，都是同人类的兴盛繁荣有关的。因此可以说"合理的"是"善"的一部分。普特南着重批判了科学主义的观点，这种观点认为人们之所以要成为合理的，是因为只有这样才能了解手段与目的之间的关系，从而有利于使手段为目的服务（"工具主义"）；是因为只有这样才能有效地发现新的真理（"方法学科"）。普特南再三强调，不仅科学领域有合理性问题（即使在这个领域中，也不能把合理性归结为一套形式化算法），而且科学以外也有合理性问题；不仅手段有合理与否的问题，目的也有合理与否的问题。

那么，到底什么是"合理性"呢？急于寻找这个答案的读者可能会失望。普特南并没有为"合理性"下一个定义，而只是在书的末尾向读者许诺一个完善的合理性观念存在于未来，而不存在于过去和现在："有没有一个真的合理性观念，一种真的道德，即使我们所曾有的都是我们对于这些东西的观念？……我们把我们的不同观

念看作是对合理性的不同观念来谈论这一事实本身，便设定了一个极限概念，一个理想真理的极限概念”。这就是说，就任何具体时刻（如现在）而言，并不存在一个本身是真的合理性观念。但是通过不断的研究、讨论和对话，我们总会得出越来越真的“合理性”观念，它们总是朝着作为极限的“真”的合理性观念会聚的。从这里我们看到，不仅真理问题的解决要取决于合理性问题的解决，而且合理性问题的解决也取决于真理问题的解决。

说真理问题的解决取决于合理性问题的解决，还需要另一个概念作为过渡。因为“真”是陈述的性质，而“合理”则是人的行为的性质。因此，普特南又用了“合理的可接受性”的概念。如果一个陈述被人们认为是有理由接受的，即接受这个陈述被人们认为是合理的，那么这个陈述就具有“合理的可接受性”。在普特南看来，如果一个陈述是真的，它一定具有合理的可接受性。由于“内在的实在论”否认了陈述同陈述系统之外的“事态”之间的符合，所以只能谈论我们的信念之间、信念与信念系统中出现的经验之间的融贯，也就是说，只能从我们认为是否有理由来接受的角度来评价一陈述或理论。但是，并不能反过来说，只要是具有合理的可接受性的，就是真的。事实是，常常一个理论可接受，却并不真。普特南说，这是因为“合理的可接受性和真理之间的关系是两个不同概念之间的关系。一个陈述可能一时是合理地可接受的，但却并不是真的”（序言，第2页）。在这方面，普特南主要反对当代最流行的两大合理性观念。一是逻辑实证主义的，一是相对主义或无政府主义科学哲学或文化哲学的。逻辑实证主义者持一种“惯例化的唯标准的合理性”概念，但这个观点本身并不符合它所规定的合理性标准（如“可证实性”）。在普特南看来，这种合理性观念的视界太窄——一方面把精密科学之外的领域都排除了，另一方面把探讨别的、多样化的合理性标准的道路堵塞了。但也不能像库恩、尤其是法伊尔阿本德那样根本否认在不同的合理性观念之间有可能进行比较、对

话和相互理解，并因此而否定合理性概念的存在。普特南认为，相对主义者之所以错误，一是因为他们显然无视科学发展中不同理论之间的“会聚”，二是因为他们无视这样一个简单的事实：我们不仅把此刻的自己，而且把先前的自己、我们的前辈、过去和现在的其他文化的成员等等都当作“人”来对待，而不是当作只会发出莫名其妙声音的动物来对待。从理论上说，其错误根源在于混淆了“概念”(concept)和“观念”(conception)。尽管现代科学家和18世纪的科学家对“温度”这个concept可以有不同的conceptions，即不同的描述、信念和理解等等，但既然拥有这同一个concept，他们就拥有相同的指称，因而现代科学家既能够理解18世纪科学家的工作，又能够比他们更接近对“温度”这个concept的真的conception。这也就是说，普特南一方而承认“合理的可接受性”标准的多样性和可变性，另一方面又不因此而否定真理的存在。作为实现这两方而观点之间综合的一个方案，普特南对真理作了这样一个规定：真是合理可接受性的一种理想化，或一种理想化的可接受性。具体地说，我们可以假定有一些“认知上的理想条件”，如果一陈述在这种条件下被确定为正当，它就是“真”的。“认知上的理想条件”就像物理学中的“无摩擦平面”一样，虽然我们永远不能真正得到，但我们可以永远向它逼近。把这个思想反过来再运用于“合理性”观念本身，就得出了前而提到的那个观点，即存在着一个作为极限概念的“合理性”，这既不同于否认有可能进行对话的法伊尔阿本德，也不同于只承认对话的罗蒂(关于罗蒂对普特南在本书末尾对他的批评的反驳，可参见已译成中文的罗蒂名著《哲学和自然之镜》最后部分)。

由于把“合理性”概念从科学领域即所谓“事实”领域推广到道德、文学、政治和历史等等领域即所谓“价值”领域，由于通过“合理性”而把“真”与“善”联系起来，显然，以往哲学家们所坚持的事实与价值之间的非此即彼关系也就被解除了。在普特南看来，所谓事

实，就是可合理地相信的东西。更确切地说，“事实”（或真陈述）的概念，就是“可合理相信的陈述”这个概念的理想化。“合理的”不仅意味着具有合理可接受性的标准，而且意味着具有相关性的标准，而我们所有的价值都包含在我们的相关性标准之中。被认为用来衡量科学陈述之客观性的那些标准，如自洽、简单等等，其实也是一些价值。另一方面，价值也不是仅同个人的趣味有关的任意的东西。对边沁派的那个著名命题——“音乐和诗的艺术或科学”同儿童的针戏游戏没有客观的高低之分，普特南认为是不对的。普特南因而强调，价值判断也是有理由的，当它们同人生有重要联系的时候，更不能强调它们的主观任意性，而不注意客观上它们有没有合理根据。在用这个观点分析福科的历史学著作时，作者反对后者关于意识形态信仰和与之相关的价值判断是一件完全任意的事情的观点，再一次强调对意识形态之类的东西，不仅要考察它所产生的非理性的根源，如马克思说的经济利益，弗洛伊德说的无意识冲动和尼采说的强力意志之类，而且要看它曾经有没有理由，现在在我们看来有没有理由。

总之，普特南的《理性，真理与历史》一书探讨了一系列重大的哲学问题，尤其是心灵和世界的关系问题。他的回答是，既非心灵单方面地创造了世界，也非世界单方面地创造了心灵，而是心灵与世界一起创造心灵与世界。从这个“内在的实在论”的立场出发，他力图克服主观与客观之间、事实与价值之间、实在论与唯心论之间、真理的符合论与真理的融贯论之间、逻辑主义与历史主义之间、独断论与相对主义之间等等的一系列对立。作者在分析批判各家哲学的弱点时所运用的严谨方法，所具有的锐利眼光，他对于当代哲学重大问题和主要潮流在把握上的透彻、广泛，在阐述上的深入浅出，对于我们是很有启发的。

但是，对于作者自己提出的建设性结论，我们不能作过高估计。对普特南的基本哲学立场，我们有理由提出批评。试举几例。

先看内在论。普特南的内在论从理论上可以追溯到洛克尤其是康德的某些思想。洛克曾作过第一性质和第二性质的区分。第二性质是主观感觉。普特南把洛克对第二性质的说法推广到第一性质上去，即不仅运用到颜色、味道等主观感觉性质上，也运用到形状、大小等外部对象的性质上，因而得出一切对象都不独立于我们的概念框架而存在，只能从人所制定的理论之内谈论对象才有意义的结论。这是一种扩展了的洛克哲学。

对康德哲学，普特南也作了自己的解释。他认为，尽管康德哲学中有自在之物和为我之物的区分，但康德并不认为为我之物（观念）和自在之物之间有对应关系。这不仅说明康德没有符合论思想，也为他自己坚持内在论提供了依据。普特南甚至认为自在之物本身是个不必要的形而上学要素。因为在他看来，对任何经验对象的谈论都不是对物自体的谈论，而仅仅是对为我之物的谈论，所以根本不可能有什么自在之物。显然这是一种主观化的康德哲学。

至于内在论和外在论的对立，也没有普特南想象的那么严格。我们很容易找到一种既非内在论，又非外在论的哲学观。这就是真理的多余论。持这种观点的人认为真理是个多余的概念，无须作进一步说明。既然他无意接受符合论，他就不是形而上学实在论；既然他不想构造真理、合理性和正当性证明等概念之间的联系，他就不是内在论者。如果这是可能的，普特南关于外在论和内在论的非此即彼的区分就下得过早了。而且和他所反对的众多的二分法一样，他自己也陷入一个新的二分法，等待别人去瓦解。

另有一些实在论者对普特南用三个命题来概括形而上学实在论表示异议。南加州大学的费尔德就认为，事实上这三个命题代表三种观点，每个之间并无必然联系。例如持世界是独立于心灵而存在的这一观点的人，并不一定坚持对世界只有一种真实而全面的

描述，也不必持有真理的符合论。[1]

至于真理是理想化的合理的可接受性，也有可疑之处，显然，有许多日常的真理通常是非理性地接受下来的，例如爱斯基摩人的雪屋能保暖等。有一些真理有时甚至是无法合理地接受的，例如存在过的恐龙的数目不会超过 $10^{10^{10}}$，结果，要么 0，要么 1，要么 $10^{10^{10}}$ 个恐龙存在过。因此，下列陈述必定有一个是真的：

S_0 存在过 0 个恐龙，

S_1 存在过 1 个恐龙，

⋮

$S10^{10^{10}}$ 存在过 $10^{10^{10}}$ 个恐龙。

如果真是合理的可接受性，这 $10^{10^{10}}$ 之一个陈述总是合理可接受的。但我们接受哪一个是合理的呢？显然，接受任何一个都没有道理。因为我们没有任何证据使得其中之一被合理地接受下来。

根据普特南的想法，只要接受经过扩展的洛克理论，即把一切性质都看作第二性质，就等于自动放弃符合论。事实上我们很容易找到反例。如，我们可以说："我眼前有一个棕色对象。"不管这个对象是一个客观对象，还是一种主观性质，都可以被符合论真理观认可为真。

普特南对各种形式的相对主义的批评也招致了不少相对主义的辩护者们的批评。费尔德认为，相对主义应作进一步区分。有认识论上的相对主义，也有价值上的相对主义。价值相对主义并不隐含真理的相对主义。费尔德进一步为认识论的相对主义作了大量辩护。按他的解释，认识论相对主义讲的是，某一信念 B 的正当性证明只是相对于证据系统 E 的，没有绝对的正当性证明的标准。同时，认识论相对主义并不宣称全部证据系统无好坏之分。因此普

---

① 参见 H·费尔德："实在论和相对主义"，载《哲学杂志》，第 79 卷，1982 年 10 月号。

特南对于相对主义的有关反驳就是无的放矢了。[1] 另一位普林斯顿大学的哈曼教授则着重为道德相对主义辩护。他认为我们可以假定有运用于全人类的合理性原则，但没有可运用于全人类的道德原则。哪怕不偷盗不说谎这类通常被认为是基本道德准则的东西，也是无法作出正当性证明的。例如，要证明不偷盗是对的，你必须用中性词来表述。但事实上人们通常只是说偷盗是错的，不应偷盗，这等于同语反复。[2]

由于普特南一方面实际上偏向于非实在论一边，另一方面又力求克服不同观点的对立，我们还可以在本书中找到不少含糊的甚至矛盾的地方。举个例子来说，普特南既承认实在，又认为只能在理论之内谈论实在。但什么叫"谈论"实在呢？是谈论实在的具体性质和关系吗？若这样的话，强调只能在理论之内谈论实在算不了太勇敢的举动。因为实在的性质和关系一定是由理论来刻划的。那么是指谈论实在（及其性质、关系）的存在？如果这样的话，问题就出现了：理论是会发展的（这点普特南承认），因而是会出现新的研究对象的。如果我们要合理地说明科学发展中的这个事实（普特南本人要求实在论所具有的正是这个说明功能），我们有没有必要认为（或至少假定）曾经有过一个在这个理论之外的对象呢？不然的话，岂不等于断言理论的存在单方面地决定了对象的存在吗？而在这种情况下，用"唯心论"这个标签来代替实在论这个标签岂不是更合适一些吗？当然，普特南会说，即使"存在"这个术语，在不同的理论中也可以有不同的意义。但这毕竟是两个层次的问题。更何况，按普特南本人的原则，对于"存在"这个"概念"的各种不同的"观念"不是也会向一个理想的、真的"存在"观念会聚吗？不管这个

---

① 参见H·费尔德："实在论和相对主义"，载《哲学杂志》，1982年10月号。

② 参见G·哈曼："形而上学实在论和道德相对主义"，载《哲学杂志》，第79卷，1982年10月号。

"真"的"存在"观念的具体内容是什么,我们难道不可以根据这样的"存在"观念抽象地说,在一个理论出现之前,它的对象已经存在着吗?

在发表《理性,真理与历史》的前后数年中,普特南发表了专著《意义和人文科学》和三卷论文集《数学,物质和方法》、《心,语言和实在》和《实在论和理性》。在这些著作之后,他又发表了《实在主义的诸多面孔》、《表征和实在》和《具有人类面孔的实在主义》等著作。在这些著作中,普特南的出发点都是《理性,真理与历史》这本书,但他不仅提出了许多新的问题和观点,对别人的批评(包括本文提到的一些)进行了答辩,而且修改、甚至放弃了自己原先提出的一些观点(尤其是科学实在论和功能主义的心灵观)。译者之一近年来对德国哲学家尤根·哈贝马斯的工作比较有兴趣,因而觉得普特南后来著作中的这个倾向特别有意思:普特南越来越多地提到哈贝马斯的观点,并非常明确表示他对他所谓的"新法兰克福学派"的观点(尤其与法国后结构主义哲学相比)的赞同。本文一开始介绍了普特南的社会政治立场和他对于哲学的看法,这种立场和看法是普特南与哈贝马斯之间共识的重要基础。但从哲学上讲,更重要的是他们从不同的**理论前提**所得出的相当接近的**理论结论**。在"解构"所谓"合理性霸权"本身几乎成为一种"霸权"的当代西方哲学界,普特南和哈贝马斯在一些重大哲学问题上的共识,尤其值得人们慎重对待、认真研究。但这需要更多的时间。

本书由童世骏译一、二、三、四、五章,李光程译序言、六、七、八、九章。译文不当之处,敬请专家、学者及广大读者批评指正。

童世骏 1988 年 9 月 16 日初稿于卑尔根

李光程 1988 年 12 月 28 日二稿于剑桥

童世骏 1994 年 12 月 12 日补充于上海

献给露丝·安娜

目　录

序 言

本书旨在打破为数众多的二分法对哲学家和非哲学家思想的束缚。在这些二分法中，最主要的要数有关真理和理性的主观论和客观论的二分法。我所关注的现象乃是：一旦承认“主观”和“客观”的二分法（不是仅仅承认为一对范畴，而是承认为对观点类型和思考风格的一种刻划），思想家们便开始将二分法的词汇视为一种意识形态标签。当今许多（或许是大多数）哲学家所持有的那种类型的“摹本”说认为，一陈述为真，仅当该陈述“符合（独立于心灵的）事实”；这一派哲学家还认为，除此之外的唯一选择，是否认真理的客观性，并且倾向把一切思想体系和观点都看作完全是主观的那种立场。而那些勇敢的少数派（例如，库恩，至少就其气质而言；法伊尔阿本德以及像福科这类杰出的大陆哲学家们）则必然将自己划归这个敌对阵营。他们一致认为，在这样一个素朴的真理摹本说之外的唯一选择，便是将思想体系、意识形态、甚至科学理论（在库恩和法伊尔阿本德那里）看作是主观的，并进而不遗余力地提出一种相对主义的和主观的观点。

哲学争论或多或少体现着意识形态争端的性质，这一点本身并不一定是坏事：哪怕在最严格的科学领域里，对新观念的拥护或攻击通常也都是带有党派热情的。即使在政治学中两极分化和意识形态热情有时也不可避免地使道德的严肃性发生疑问。但是，无论是哲学上还是政治上，新观念终究会成为旧观念：那曾具有挑战意义的观念、会成为陈词滥调，令人生厌；那曾一度在理应引人注

目之处引人注目的东西，随后却会成为新的选择的绊脚石。这种情形现在正出现在种种真理的符合论和种种主观主义真理观的争论之中。在本书的前三章，我试图阐明一种将客观论和主观论的成分融为一体的真理观。这种观点，至少在精神上回到了康德的思想，即是说，我们无须宣称真理只是某种时代精神，或者真理只是某种格式塔转换，或者真理只是某种意识形态，便可拒绝素朴的真理“摹本”说。

我要辩护的观点，简单地说是，在真理概念和合理性概念之间有着极其密切的联系。粗略说来，用以判断什么是事实的唯一标准就是什么能合理地加以接受。（我的意思确实如此，而且适用于一切对象，因此，如果一幅画是美丽的这一点能合理地接受，则这幅画是美丽的这一点即可成为事实。）根据这个观点，可以有价值事实存在。但是，合理的可接受性和真理之间的关系是两个不同概念之间的关系。一个陈述可能一时是合理地可接受的，但却并不是真的。在我对合理性概念的说明中，我将为这个实在论的直觉保留一席之地。

但是，对合理性能被一系列不变的“标准”或“原则”所定义这一点，我是持怀疑态度的。方法论原则与我们对世界的看法，包括对作为世界之一员的自我的看法，是相互联系的，而且是随时间而变化的。因此，我赞同主观论哲学家们的观点，用来定义合理的东西的、一成不变的、与历史无关的方法论原则是根本不存在的。但是我不想从我们对理性的看法是在历史中发展的这一事实得出结论说，理性本身就可以是（或变成是）任何东西，我也不会像一些法国哲学家那样以某种文化相对主义和“结构主义”的奇特混合体来作为我的理论归宿。要么非历史的、一成不变的合理性标准，要么文化相对主义，这种二分法我认为业已过时。

我的观点的另一特征是，合理性不局限于实验科学，并且就其基本方面而言，它在不在实验科学之中都没有区别。那种把合理性

只局限于实验科学的观念，我以为是实证主义观念的残余。实证主义观念认为，从某种程度上讲，科学世界由“感觉材料”所构成，并且实验科学的术语是被人“从操作上进行定义”的。在过去20多年中，操作主义和实证主义的科学观遭到的批判已经够深够透，因此我不打算花大量篇幅去批判它们。但是，我将根据我们在真理和合理性方面必须表达的观点，对经验主义关于“感觉材料”至少为我们知识的一部分构成某些客观“基层”的观点加以重新考察（在第三章）。

简言之，我将提出这样一种观点：心灵并不简单地“摹写”一个可由一个唯一真理论所描述的世界。当然，我的意思也不是说心灵构造了世界（或使世界构造屈从于“方法论原则”和独立于心灵的“感觉材料”所强加的制约）。如果非得使用隐喻来表达不可，那么这个隐喻可以这样说：心灵和世界一起构成心灵和世界（或者，让这个隐喻更黑格尔化一些，宇宙——和在构造中起着特殊作用的心灵一起，集体地——构造着宇宙）。

我对合理性的说明的最后一个特征是，我将试图表明，我们的合理性概念从根本上说，只是我们关于人类兴盛的观念以及善的理念的一部分。真理极大地依赖于眼下被称之为“价值”的东西（第六章）。我上面对合理性和历史的说法也同样适用于价值和历史。不存在一套一劳永逸地定义人类兴盛之内容的、既定的、非历史的道德原则。但这并不意味着这一切仅仅是唯文化的和相对的。在我看来，既然当前真理论的现状——摹本说真理观和主观论真理观在当前的二分法——和众所周知的事实/价值的二分法多少有点关联，那么只有通过推进到相当深入的水平，并且修正我们对真理和合理性本身的解释，我们才能超越事实和价值的二分法。（这个二分法，正如通常所理解的，实际上使人倒向了某种相对主义。）当前的种种真理观已经异化。它们使人失去其自身和世界的这个或那个部分，教人把世界仅仅看作是由在虚空中旋转的基本微粒

所构成(即"物理主义"的观点,这新观点把科学的描述看作是向一个唯一真理论的会聚),或者把世界看作仅仅是由"实际的和可能的感觉材料"所构成(较早的经验主义观点),或者根本否认有一个世界,而与我们出于各种各样(主要是无意识的)理由所构造的形形色色的一系列说法恰恰相反。在本书中,我的目的是勾勒一个非异化的真理观的基本轮廓。

我在赫伯特·斯宾塞讲座上的演讲"哲学家和人类理智"(作于1979年,牛津大学)和我的论文"'上帝死了,一切都是可能的'……(语言哲学的反思)"(载于《批判》,1980年)一样,与本书有部分重合,都是由于边研究边写作的缘故。

从1978年到1980年,国家科学基金会提供了一笔研究基金,[1] 资助了与本书有关的研究,在此我深表谢意。

托马斯·库恩和露丝·安娜·普特南研究了本书的手稿并提出了中肯的批评和明智的建议。我也从许多朋友的批评和建议中得到诸多帮助。他们是:内德·布洛克、大卫·赫尔曼、贾斯廷·莱伯以及我在哈佛的各种研究班和讲座的学生们。其中几章曾于1980年春在利马的演讲中宣读过(承富布赖特基金会慷慨相助,使这次利马之旅得以成行)。第二章实际上在利马期间即告完成。在此期间,我与利奥波多·奇亚波、艾伯托·科德罗·莱卡、亨里克斯·费尔南德斯、弗兰西斯科·米诺·奎萨达和乔奇·塞卡达等人就有关内容展开过讨论,从中获益匪浅。全书(和这个版本有所不同)曾于1980年夏天在法兰克福大学的讲座中宣读。我很感谢那里的同事(特别是威廉·埃斯勒和雷纳·特拉普),感谢我的极富启发性的学生和我在德国的其他朋友(特别是迪特尔·亨里奇、曼隆·法斯宾德和沃尔夫冈·斯泰格缪勒)对我的鼓励,以及

---

① 其目的是研究"科学理论的评价:历史的研究方法与形式的研究方法之对立";合同编号为SOC78-04276。

富有启发性的讨论。

我在哈佛哲学系的所有同事都值得一一致谢。近年来纳尔森·古德曼和我已经发现了我们观点上的一种趋同。虽然在本书的初稿完成之前，我无缘读到他的《世界的构成方式》一书，但后来阅读此书并与他就这些问题进行讨论，这些在本书写作过程的不少阶段上都很有价值。

最后，我还要感谢杰里米·迈诺特以编辑身分给予我的鼓励和建议。

第一章　钵中之脑

沙地上爬着一只蚂蚁。爬着爬着，蚂蚁在沙地上划出一条线来。纯粹出于偶然，这线条弯弯曲曲，最后看上去就好像是一幅清晰可辨的温斯顿·丘吉尔漫画。这蚂蚁画了一幅温斯顿·丘吉尔的像没有？也就是说，这蚂蚁画了一幅描绘丘吉尔的像没有？

多数人在稍作思考以后会说：没有。这蚂蚁毕竟从未见过丘吉尔，连丘吉尔的照片也没见过，而且，它也没有描绘丘吉尔的意向。它仅仅是划了一条线（即便这，也是无意的）、一条我们能"看作"丘吉尔之像的线而已。

我们可以这样来表达上述看法：那条线"本身"并不表征[1] 某个确定的东西。与温斯顿·丘吉尔（某种颇为复杂的）外貌特征相似，并不足以使一样东西表征或指称丘吉尔。它也不是表征或指称丘吉尔的必要条件：在我们的共同体中，"温斯顿·丘吉尔"这些字形，"温斯顿·丘吉尔"这些语词，以及其他许多东西，都被用来表征丘吉尔（虽然不是用图形），然而这些都没有一幅图画——即便

[1] 本书中"表征"(representation)和"指称"(reference)这两个术语始终用来表示一个词（或其他指号、符号或表征）同某个实际存在的东西（也就是说，不仅仅是"思想的对象"）之间的一种关系。按"指称"的一种含义，我可以"指称"不存在的东西；但这里"指称"一词不是在这种意义上使用的。我称之为"表征"或"指称"的以前有一个词叫"指示"(denotation)。

其次，我依照现代逻辑学家的惯例，用"存在"来表示"存在于过去、现在和未来"。因此，温斯顿·丘吉尔"存在"着，而我们也可以"指称"或"表征"温斯顿·丘吉尔，尽管他已经去世。

是一幅线条画——与丘吉尔之间所具有的那种相似性。如果相似性对于使某物表征别物来说既非必要也非充分，那么，为了达到表征的目的，还有什么东西可以是必要的或充分的呢？到底一个东西怎样才能表征（或"代表"，等等）另外一个东西呢？

问题看上去不难回答。假定该蚂蚁见过温斯顿·丘吉尔，假定它具有为丘吉尔画像的智力与技巧，假定它是有意画出那幅漫画的，如果那样的话，那线条就表征了丘吉尔。

相反，假定线条具有"温斯顿·丘吉尔"这几个字的笔划形状；又假定，这完全是出于偶然（其不可能性撇开不谈），那么，"温斯顿·丘吉尔"这些字形也并不表征丘吉尔，尽管这些字形出现在今天几乎任何一本书里都表征着丘吉尔。

因此，表征的必要条件，或首要条件，看来就是意向。

但是，要具有用任何东西、哪怕是私人语言（哪怕在我心中默默无声地念出的"温斯顿·丘吉尔"这几个词）来表征丘吉尔的意向，我先得能够想到丘吉尔。如果沙上线条、声音等等"本身"无法表征任何东西，那思想形式"本身"怎么就能表征任何东西呢？它们能不能呢？思想怎么能伸出手来"把握"外部事物呢？

以前有些哲学家，就是从这种考虑跳至他们所认为的对心灵本质上是非物理之物的一种证明的。论证很简单：我们关于蚂蚁划出的曲线的一番道理适用于任何物理对象。任何物理对象本身都无法指称此物而不指称彼物；然而，心中的思想显然确实能够指称此物而不指称彼物。因此，思想（从而心灵）便具有与物理对象根本不同的本质。思想具有意向性的特征——它们能指称别物；物理的东西都没有"意向性"，尽管那种意向性要通过心灵对于那种物理东西的运用而得到。至少人们是这样宣称的。这个结论来得过于仓促了；只是设定心灵的神秘力量，是解决不了任何问题的。但问题还是非常现实的，意向性、指称何以可能呢？

神秘的指称理论

我们已经看到，那只蚂蚁的“图像”同温斯顿·丘吉尔之间并没有必然联系。此“图像”与丘吉尔“相像”这个事实本身并不使之成为一幅真正的图像，也不使之成为丘吉尔的表征。除非这只蚂蚁具有智慧（但它并不具有），并且知道丘吉尔（但它并不知道），否则它所划出的曲线就不是任何东西的图像，甚至也不是任何东西的表征。有些原始人相信，有些表征（尤其是名称）同它们所表征的东西之间有一种必然的联系：知道了某人或某物的“真实名称”，就获得了征服他（它）的力量。这种力量来自名称与名称的承担者之间的神秘联系；一旦人们发现名称同它的承担者之间只有一种依赖于上下文的、偶然的和约定的联系，就难以理解为什么关于名称的知识居然会有神秘的意义。

重要的是要看到，适用于物理图像的也适用于心理图像，并且适用于一般的心理表征；就像物理表征与它们所表征的东西之间不存在必然联系一样，心理表征与它们所表征的东西之间，也没有必然的联系。与之相反的论断是神秘思维的残存物。

也许这个观点在心理意象的例子中最容易把握。（最早领会这个观点的巨大意义的哲学家大概是维特根斯坦，即使他不是最早提出这个观点的人）假定某处有一颗行星进化出了人类（或由外空人置放了人类，或以随便什么方式出现了人类）。假定这些人虽然与我们相像，却从未见过树，假定他们从未想象过树（他们星球上的植物或许只以霉菌的形式出现）。假定有一天，一艘宇宙飞船飞经他们的行星，偶然地把一幅树的图像落在这个星球上，而没有同他们发生其他联系。请设想一下他们对这幅图像所感到的困惑。这到底是什么？他们作了各种各样猜测：一幢房子，一顶天篷，甚至是某种动物。但假定他们从未猜对过。

对我们来说，这图像是树的一种表征。对于那些人来说，这图像只表征一个陌生物体，一个不知其本质和功能的陌生物体。假定他们中的一个人看了图像以后有了一种同我关于树的心理意象非常相像的心理意象。他的心理意象并不是一种对树的象征。它只是对那幅神秘图像所表征的陌生物体（不管它是什么）的一种表征。

然而，有人会坚持说，这个心理意象实际上是树的表征，只要引起这个心理意象的图像本身先是树的表征。从实际的树到这个心理意向之间有一条因果之链，尽管这是一条非常奇特的因果之链。

但是，即使这条因果之链也可以设想成并不存在，假定那艘宇宙飞船掉下来的“树的图像”实际上并不是树的图像，而是一些溅出的颜料的偶然结果。即使它看上去同树的图像非常相像，但它实际上并不是树的图像，就像那只蚂蚁的丘吉尔“漫画”并不是丘吉尔的图像一样。我们甚至可以设想那艘掉下此图像的宇宙飞船来自一颗对树一无所知的行星。这样，那些人仍然会具有同我对树的意象性质上相同的心理意象，然而它们并不可能是表征一棵树的意象，就像不可能是表征任何别的东西的意象一样。

这同样也适用于语词。纸上的一席话可能会显得是对树的确切描绘，但是假如它是由猴子们在几百万年间随机地敲打打字机的键盘而产生出来的，那么这些语词也并不指称任何东西。假如有一个人记住这些语词、把它们默诵出来，但没有理解这些语词，那么，它们被人在心中思考时也不指称任何东西。

试设想，默诵这些语词的人已经被催眠了。假定他说的是日语，而且他被告知他懂日语。假定当他思考这些语词时他具有一种“理解感”。（虽然如果有人打断他的思路并问他，他正思考的话的意思是什么，他会发现自己无言以对。）或许这人的幻觉可以逼真得连一位日本传心术士也遭愚弄。但是，假如他不能在恰当的上下

文中使用那些语词，并回答有关他的"思考"的问题，等等，那么，他对这些语词就并不理解。

把我讲的这些科学幻想故事联系起来，我们就能设想出这样一种情况，即某人思考一些在某种语言中确实是对树的描述的语词，并且同时具有合适的心理意象，然而既不理解这些语词，也不知道树是什么。我们甚至可以设想那种心理意象是由颜料四溅所引起的（尽管此人经过催眠后以为它们是适合他思考的某物的意象——只不过，如果问他，他会说不出到底思考什么）。我们还可以设想，此人用来思考的语言是催眠师和被催眠者都从未听到过的——或许，这些被催眠师看作"胡说八道"的句子是用日语对树的描绘这一点，只是一种巧合。简言之，那个人想到的一切可能同一个真正思考树的说日语的人所想到的东西具有相同性质——但它们没有指称树。

当然，所有这些实际上都是不可能的，就像猴子碰巧打字打出一本《哈姆雷特》来实际上不可能一样。也就是说，与之相反的概率之高，使得它实际上决不会发生（我们是这样认为的）。但它并不是逻辑上不可能的，甚至也不是物理上不可能的。它有可能发生（它同物理定律相容，或许也同宇宙间的现实条件相容，如果在其他行星上存在许多智慧生物的话）。如果这种情况真的发生，它就会以引人注目的形式证明一条重要的概念真理：即使一个大而复杂的表征系统（既包括言语的，也包括视觉的），与它所表征的东西之间也不具有内在的、固有的、神秘的联系——一种与它被怎样引起的过程无关、与说话者和思想者有什么行为倾向无关的联系。不管表征（在我们举的例子中，即语词和意象）系统具有物理的实在——语词被写下或说出，图像是物理的图像——还是只具有心理的实在，上述观点都是对的。想到的语词和心理图像并不内在地表征它们与之有关的东西。

以钵中之脑为例

这里有一个哲学家们所讨论的科学幻想中的可能事件：设想一个人（你可以设想这正是阁下本人）被一位邪恶的科学家作了一次手术。此人的大脑（阁下的大脑）被从身体上截下并放入一个营养钵，以使之存活。神经末梢同一台超科学的计算机相连接，这台计算机使这个大脑的主人具有一切如常的幻觉。人群，物体，天空，等等，似乎都存在着，但实际上此人（即阁下）所经验到的一切都是从那架计算机传输到神经末梢的电子脉冲的结果。这台计算机十分聪明，此人若要抬起手来，计算机发出的反馈就会使他“看到”并“感到”手正被抬起。不仅如此，那位邪恶的科学家还可以通过变换程序使得受害者“经验到”（即幻觉到）这个邪恶科学家所希望的任何情境或环境。他还可以消除脑手术的痕迹，从而该受害者将觉得自己一直是处于这种环境的。这位受害者甚至还会以为他正坐着读书，读的就是这样一个有趣但荒唐之极的假定：一个邪恶的科学家把人脑从人体上截下并放入营养钵中使之存活。神经末梢据说接上了一台超科学的计算机，它使这个大脑的主人具有如此这般的幻觉……

如果在一次关于知识论的讲演中提到这种可能事件，目的当然是要以现代的方式提出那个关于外部世界的经典的怀疑论问题。（你怎么知道你现在不处在这种困境之中？）但是这种困境对于提出有关心灵与世界的关系问题，也是一个有用的手段。

我们可以设想，不只是一个大脑放在钵中，相反，所有人类（或许所有有感觉的生物）之脑都在钵中（如果某些只具有最低级神经系统的生物也已经算作“有感觉”，那就是钵里的神经系统）。当然，那个邪恶的科学家必须在营养钵之外——要不他愿意吗？或许并没有邪恶的科学家，或许这宇宙恰好就是管理一只充满大脑和神

经系统的营养钵的一台自动机(尽管这是荒谬的)。

这会儿让我们假定这架自动机具有这样的程序,它向我们大家提供一种集体幻觉,而不是若干互不相干的幻觉。这样,当我觉得自己正与你交谈时,你则觉得正听我讲话,当然,我的话并没有真的进入你的耳朵——因为你并没有(真实的)耳朵,我也没有真实的嘴巴和舌头。相反,当我讲话时,所发生的是外输脉冲从我的大脑传到计算机,该计算机既引起我"听到"我讲这些话的声音和"感到"我的舌头颤动等等,也引起你"听到"我的话,"看到"我在讲话,等等。在这种情况下,我们在某种意义上真的进行着交流。你实际存在着,对此我并没有搞错(我搞错的只是除了大脑,还存在着你的身体和"外部世界")。从某种观点来看,即使"整个世界"是一个集体幻觉,也无碍大局;因为当我对你讲话时,你毕竟确实听到了我的话,即使起作用的机制并不如我们所设想的那样。(当然,如果我们是正在做爱的两个情人,而不只是正在交谈的两个人,那么,说这只是一个钵中的两个大脑,会带来不少麻烦。)

现在我想提出一个看上去非常愚蠢、又非常浅显的问题(至少对有些人,包括有些相当老练的哲学家),但它将使我们很快达到真正的哲学的深度。假定上述全部说法确实是真的,如我们以这种方式成了一个钵中之脑,我们能不能说或想到我们是钵中之脑?

我将要论证,对这个问题的回答是"不,我们不能"。事实上,我要论证,说我们真的是钵中之脑这样的假设,虽然并不违反物理定律,并且同我们的所有经验完全一致,但却不可能是真实的。它不可能真实,因为它在某种程度上是自相反驳的。

我将要提出的论证不同寻常,为了使自己确信这个论证确实无误,我花了几年时间。但这是一个正确的论证。它之所以显得奇怪,是因为它同某些最为深刻的哲学问题有关。(我是在考虑现代逻辑中的一个定理即司寇伦-勒文海姆定理的时候,首次想到这个论证的。我突然看到了这个定理与维特根斯坦《哲学研究》中某些

论证之间的某种联系。）

“自相反驳的假设”是这样的假设，它的真蕴涵着它自己的假。例如，试考虑“凡全称命题皆假”这个论题。这是一个全称命题。所以，假如这个命题是真的，它就一定是假的。所以，它是假的。有时，只要一个论题假设它包含或表达着蕴涵它的假，该论题便可被称为“自我反驳”。例如，“我不存在”这个命题如果是我所想到的，它就是自我反驳的（对于任何“我”来说都是这样）。所以，如果某人想到这个命题，他就可以确信他自己存在着（就像笛卡儿证明的那样）。

我下面要表明的就是，那个说我们是钵中之脑的假设正具有这个性质。如果我们能够考虑它是真的还是假的，它就不是真的（我将表明这一点）。因此，它不是真的。

在作出这个论证之前，先让我们考虑一下，为什么能够作出这样一个论证这一点显得那么奇怪（至少对那些同意把真理看作“摹本”的观点的哲学家来说）。我们承认，可能有一个使得一切有感觉的生物都是钵中之脑的世界，这是符合物理规律的。如同哲学家们所说的，有一个“可能世界”，其中一切有感觉的生物是钵中之脑。（这个“可能世界”的说法听起来好像是，有一个地方，在那里任何荒谬的假定都是真的，这就是它在哲学上会产生严重误解的原因。）那个可能世界中的人具有与我们完全相同的经验，具有与我们完全相同的思想（至少，他们从中发现同样的语词、意象、思想形式等等）。但是，我要说的是，我们可以给出一个论证，表明我们并不是钵中之脑。如何能有这个论证呢？为什么真的是钵中之脑的可能世界中的人们不能够也提出这个论证呢？

回答将（基本上）是这样的：可能世界中的人们虽然能够想和“说”我们所能想、所能说的任何话语，但（我认为）他们不能指称我们所指称的东西。尤其是，他们无法想或说他们是钵中之脑（即使通过心想“我们是钵中之脑”）。

图灵测验

假定某人成功地发明了这样一架计算机，它真能同某人进行一场富有智慧的交谈（其话题之多不亚于一个有智慧的人）。怎么来断定这计算机有没有“意识”呢？

英国逻辑学家艾伦·图灵[①] 提出下列测验：让某人同计算机进行交谈并且同一个他不认识的人进行交谈。如果此人无法分辨出哪个是计算机，哪个是人，那么（假定此测验由不同的对话人重复足够多的次数），这计算机就是有意识的。简言之，如果计算机能通过“图灵测验”，它就是有意识的。（当然，交谈不能面对面进行，因为那位对话人不可以知道他的两位交谈伙伴中任何一位的视觉形象。也不能用声音交谈，因为机器的声音与人的声音听起来不同。相反，试设想交谈全是通过电子打字机进行的。对话人打出他们的陈述、问题等等，两位伙伴——机器和人——也通过电子键盘作出反应，而且，机器还会撒谎——当问它“你是机器吗？”时，它会答道：“不，我是这个实验室的助理。”）

以为这种测验确实是关于意识的判决性测验的观点，受到了大量作者的批评（他们决不是原则上敌视机器可以有意识的观点）。但我们眼下不讨论这个问题。我想借用图灵测验的一般观点，即对话式能力测验的一般观点，来达到另一个目的，即用来考察指称这个概念。

设想这样一个情境，其中的问题不是要判定伙伴是人还是机器，而是要判定该伙伴是不是像我们一样用语词来指称对象。这个假设的测验很明白，也是进行一场交谈，而且，如果没有出现问题，

---

① A·M·图灵：“计算机器与智慧”，载《心》(1950年)，重印于A·R·安德森主编的《心与机器》。

如果对话的伙伴"通过"了测验，也就是说，如果与一个事先已经确定是说相同语言、指称某些通常对象等的人难分彼此，那么，就判定这个对话伙伴确实像我们一样指称对象。如果图灵测验为上述目的进行，也就是说，为判定存在着（共同的）指称，我将称此测验为图灵指称测验。而且，正如哲学家们曾讨论原来那个图灵测验是不是对于意识的判决性测验，也就是讨论一台不只是一次、而是常规性地"通过"该测验的机器是不是必定有意识一样，我想也以同样的方式来讨论一下这个问题：上述图灵指称测验是不是对于共同指称的判决性测验。

我们的答案将是否定的。图灵指称测验不是判决性的。它在实践上当然是一个出色的测验，但是，某人通过图灵指称测验但并不指称任何东西，这在逻辑上并非不可能（尽管其几率无疑是极小的）。我们将看到，这意味着我们可以推广我们关于语词（以及整个文本和话语）与其所指并无必然联系的观点。即使我们考虑的不是语词本身，而是决定在某些上下文中应使用哪些语词才算恰当的规则——用计算机行话来说，即使我们考虑的是用词程序——那么，除非这些程序本身指称某种超语言对象，否则还是不存在这些语词所拥有的特定指称。这一步骤对于达到下述结论具有关键意义，即作为钵中之脑的人们，根本不可能指称任何外部事物（因而不可能说什么他们是作为钵中之脑的人们）。

比方说，假定我现在正处于图灵情境（用图灵的术语来说，正在进行"模拟游戏"）之中，而我的伙伴真的是一台机器。假定该机器能够赢得这场游戏（即"通过"测验）。假定该机器装有这样的程序，使之对用英语表达的陈述、问题、评论等等能用英语作出漂亮的对答，但它并没有感觉器官（至多只有与我的电子打字机相耦合的电路），也没有运动器官（至多只有电子打字机）。（就我所知，图灵并不把具有感觉器官或运动器官作为意识或智慧的必要条件。）假设这机器不仅缺少电子眼和电子耳等等，而且也没有这样的程

序馈饲，用以进行模拟游戏、接受来自这类感觉器官的输入或控制身体。对这样一台机器我们该说什么呢？

在我看来，很明显，我们不能、也不该说这种装置有所指称。诚然，该机器对于比方说新英格兰的风光能谈得娓娓动听，但是它不能认出一棵苹果树或一只苹果、一座山或一头牛，一片田野或一个塔顶，哪怕它们就在眼前。

这只是一台造出句子来应付句子的装置。但其中没有一个句子同实在世界有任何联系。假如把两台这样的机器耦合起来，让它们彼此间进行模拟游戏，那么即使世界的其余部分全都消失，它们也会一直不断地互相"愚弄"下去。我们没有理由认为蚂蚁的"绘画"指称着温斯顿·丘吉尔，同样也没有理由认为机器对苹果的谈论指称着实在世界的苹果。

这里，之所以会有指称、意义、智慧等等的错觉，是由于我们作了这样一个关于表征的约定，根据这个约定，机器的话语指称着苹果、塔顶、新英格兰，等等。同样，之所以会有蚂蚁为丘吉尔作了漫画的错觉，也是出于这个理由。但我们是能够感知并触摸、处理苹果和田野的。我们对苹果和田野的谈论与我们同苹果和田野的非言语相互作用之间有密切联系。有若干"语言输入规则"，它使我们从对于苹果的经验出发说出"我看到一个苹果"之类的话来；还有若干"语言输出规则"，它使我们从用语言形式表达的决定（"我要去买些苹果"）出发作出言说以外的行为。语言输入规则和语言输出规则这两者缺一不可，否则就有理由认为机器的交谈（或两台机器的交谈，如果是在我们设想的两台机器互相进行模拟游戏的情况下）仅仅是一种句法游戏。当然，这是与有智慧的话语相似的句法游戏，但这相似只是如蚂蚁的曲线与一幅讽刺漫画之间的相似一样（仅此而已）。

在蚂蚁的例子中，我们还可以提出，即使温斯顿·丘吉尔从来没有存在过，这蚂蚁也会画同样的曲线。但在机器的例子中，我们

却不能够提出同类的论证；假如苹果、树木、塔顶和田野不曾存在，那么程序编制者大概就不会编制出同样的程序。尽管机器并不感知到苹果、田野或塔顶，但它的设计制造者是感知到的。借助设计制造者的感知经验和知识的中介，在机器与实在世界的苹果等等之间有某种因果联系。但这样一种微弱的联系很难成为指称的充分条件。不仅在逻辑上可以设想下列情况（虽然它的几率小得难以想象），即哪怕苹果、田野和塔顶不曾存在，这机器也可能存在；更重要的是，这机器对于苹果、田野和塔顶等等的继续存在是完全无动于衷的。即使这些东西都不再存在，那机器仍然会一如既往地快活地交谈。这就是不能认为机器有什么指称的理由。

与我们的讨论有关的要点在于，图灵测验并不排除这样一台机器，它的程序使它只进行模拟游戏，而只做模拟游戏的机器显然不指称什么东西，就像一台电唱机那样。

再论钵中之脑

我们把假定的"钵中之脑"和上述机器作一比较，其中显然有一些重要区别。钵中之脑虽然没有感觉器官，但它们确实具有给予感官的馈饲。也就是说，有传入神经末梢，有来自这些传入神经末梢的输入信号，而且这些输入信号在钵中之脑的"程序"中的运算与它们在我们大脑程序中的运算并无二致。钵中之脑毕竟是大脑；而且，它们还是发挥功能的大脑，而它们发挥功能时与实在世界中的大脑遵循同样的规则。出于这些理由，不承认它们具有意识或智慧是荒谬的。但它们有意识、有智慧这一点并不意味着它们的语词指称着我们的语词所指称的东西。我们感兴趣的问题在于：它们的那些包含比方说"树"这个语词的言语表达，是不是真的指称树？更一般地说，它们到底能不能指称外部对象？（这外部对象与某种对象，比方说自动机产生的意象中的对象不同。）

为确定我的观点，我们来作这样的一个规定：假定这自动机是由某种宇宙机遇或巧合产生的（要不，或许是从来就有的）。在这个假想的世界中，自动机本身不应该有具有智慧的设计制造者。实际上，正如我们在本章开头所讲的，我们甚至可以设想全部有感觉的生物（不管其感觉能力多么低下）都置身于一个营养钵里。

这个假定无济于事。因为在那些大脑所使用的"树"这个语词同现实的树之间并无联系。即使不存在现实的树，它们仍将会一如既往地行事，该用"树"这个语词，就用这个语词，该有什么思想和意念，就有什么思想和意念。它们的意象、语词等等与确实表征我们世界中的树的意象、语词等等具有相同的性质；但我们已经看到（还是那个蚂蚁的例子！），一物同表征一对象（温斯顿·丘吉尔或树）的某物的性质相似并不使该物独自成为一种表征。简言之，那些钵中之脑在想到"我面前有一棵树"时并没有想到实在的树，因为并没有什么东西使它们的思想之"树"得以表征现实之树。

如果这个结论显得过于突然，那么请考虑这点：我们已经看到，语词即使被安排在一个同无疑是关于现实世界之树的话语（如果它出现在我们当中某一位心里的话）相同的系列中，也不一定指称树。规则、惯例、大脑的言语行为的倾向这样的"程序"也并不因为它在语词与语词之间、或语言提示与语言对答之间建立了联系，而必然指称树或导致对树的指称。如果这些大脑考虑、指称、表征树（实在的、营养钵之外的树），那一定是因为"程序"把语言系统同非言语的输入与输出联系了起来：在钵中之脑的世界里确实也有这种非言语输入与输出（又是这些传入与传出神经末梢！），但是我们又知道，自动机所产生的"感觉材料"即使同我们的树的意象非常相像，也不表征树（或任何外部事物）。正如颜料的溅出可以同树的图像相像而并非树的图像，我们看到，一种"感觉材料"也可以同"树的意象"性质上相同而并非树的意象。在钵中之脑的例子中，程序把语言同那些并没有内在地或外在地表征树（或任何外部事物）

的感觉输入联系起来，但这个事实怎么可能使得整个表征系统、即使用着的语言真的指称或表征树或任何外部的事物？

回答是：不可能。整个感觉材料的系统，发给传出末梢的活动信号，以及由"语言输入规则"同作为输入信号的感觉材料（或随便什么东西）相联系的、由"语言输出规则"同作为输出信号的活动信号相联系的思想，不管是以言语为中介的，还是以概念为中介的，同树都没有什么联系，就好像蚂蚁的曲线同温斯顿·丘吉尔没有什么联系一样。一旦我们看到钵中之脑的思想与现实世界中某人的思想之间的性质相似（如果你愿意，也可以说性质同一）并不意味着指称的等同，就不难看出，认为钵中之脑指称外部事物，是完全没有根据的。

此论证的几个前提

至此我已经给出了前面许诺提供的论证，即钵中之脑不可能想到或说出它们是钵中之脑。剩下要做的只是使这个论证更明白，并考察它的结构。

根据刚才所说的观点，（在那个其中任何有感觉的生物都是并且永远是钵中之脑的世界里）钵中之脑在想到"我面前有一棵树"时，他的思想并不指称现实之树。根据我们将要讨论的一些理论，它可能指称意象之树，或指称引起树的经验的电子脉冲，或指称引起那些电子脉冲的程序的特征。这些理论同刚才所说的观点并不冲突，因为在钵中-英语里"树"这个语词的使用同意象中树的出现、某种电子脉冲的出现以及机器程序中某些结构的出现之间，有一种紧密的因果联系。根据这些理论，那个钵中之脑想到"我面前有一棵树"时是对的，并没有错。如果在钵中-英语里"树"的所指已定，"在面前"的所指已定，假设上述一种理论正确，那么"我面前有一棵树"出现于钵中-英语时的真值条件仅在于意象之树在钵中之

“我”的“面前”——在意象中，或者有可能的话，在于那种通常产生这样的经验的电子脉冲是发自于自动机，或者，在于应产生出“某人面前的树”这个经验的机器结构正起着作用。这些真值条件当然被满足了。

基于同样的论据，在钵中-英语里“钵”指称意象之钵，或有关的什么东西（电子脉冲或程序结构），但肯定不是指称实在之钵，因为在钵中-英语里“钵”的使用同实在之钵没有因果联系（没有这种联系，钵中之脑就无法使用“钵”这个语词，如果这个语词不是用于一个特定的钵即它们所处的那个钵的话；但是，钵中-英语里每个语词的使用同那个特定钵之间的联系还是存在的；这不是在“钵”这个特定的语词的使用同钵之间的特殊联系）。同样，钵中-英语里“营养液”指谓一种意象中的液体，或有关的什么东西（电子脉冲或程序结构）。由此可知，如果它们的“可能世界”实际上是现实世界，而我们实际上是钵中之脑，那么我们现在说“我们是钵中之脑”这句话的意思便是我们是意象中的钵中之脑或诸如此类的东西（如果我们毕竟表达了什么意思的话）。但是，说我们是钵中之脑，这个假说部分内容是说我们不是意象中的钵中之脑（也就是说，我们所“幻觉的”不是我们是钵中之脑）。所以，如果我们是钵中之脑，那么“我们是钵中之脑”就是一句假话（如果它说了什么的话）。简言之，如果我们是钵中之脑，那么“我们是钵中之脑”就是假的，因此该假设（必定）是假的。

之所以假定钵中之脑是可能的，是有意义的，在于同时犯了这两个错误：(1)过于看重物理的可能性；以及(2)不知不觉地运用一种神秘的指称理论，根据这种理论，某些心理表征必定指称某些外部事物或某几类外部事物。

我们作为钵中之脑而置身其中的“物理可能世界”是存在的——这除了意味着对这样一种与物理学规律相容的事态的描述是存在的之外，还能作何解释？在我们的文化中（自17世纪以来），有

一种把物理学当作我们的形而上学、即把各门精密科学看作对于"宇宙之真实而终极的底蕴"的寻求已久的描绘的倾向，同样，作为这种倾向的直接后果，也有一种把"物理可能性"当作什么是实际上真会出现的事情的试金石的倾向。根据这种观点，真理就是物理学真理；可能性就是物理的可能性；而必然性就是物理的必然性。但我们已经看到，只要涉及一个精心设置的事例，这个观点就是错的。存在一个"物理可能世界"，在其中我们都是(过去一直是而且将来永远是)钵中之脑，这并不意味着我们真有可能成为实际的钵中之脑。排除这种可能性的不是物理学，而是哲学。

有些哲学家对于他们的同行的主张同时持有两种态度，既加以肯定又表示轻视(这是20世纪英美哲学的典型心态)，他们会说："没错。你表明的就是有些看起来物理上可能的事情实际上在概念上是不可能的，那又有什么奇怪呢？"

好，就算我的论证可以叫做一种"概念上的"论证，但是，把哲学活动描绘为对于概念"真理"的寻求，使之听起来完全像是对语词意义的探究，这根本不是我们一直在从事的工作。

我们所从事的工作是考察一下想到、表征、指称这些活动的前提。我们对这些前提的研究靠的不是对这些语词和短语的意义的研究(如果是语言学家的话，他会这样做)，而是先验推理。这里，先验推理并没有过去所谓"绝对"的性质(因为我们并不断言神秘的指称论是先验地错的)，而是先假定某些一般前提，或做出某些非常宽泛的理论假设，然后去考察什么样的可能性是合乎情理的。这样一种程序既不是"经验的"，也不完全是"先验的"，而是这两种研究方法兼而有之。我的程序是可错的，依赖于可称为"经验的"假设(如这样一个假定：心灵若不经过感觉的中介就无法接触外部事物或属性)，尽管如此，我的程序同康德所谓的"先验"研究有密切联系；因为，我再说一遍，它是一种对于指称、从而也是思想的先决条件的研究，这种先决条件是我们的心灵自身的本质所固有的，尽管

不是(如康德所希望的)完全独立于经验假设的。

这个论证的前提之一很明白：神秘的指称论是错的，不仅对于物理的表征是错的，而且对于心理的表征也是错的。另一个前提是，如果人们与某些事物(比方说树)[1]根本没有因果相互作用，或者与可以用来描绘它们的东西根本没有因果联系，那就不可能去指称它们。但是为什么我们要接受这些前提呢？因为这些前提构成了我在其中论证的大框架，现在该对它们作一个更仔细的考察了。

否定表征与其所指间有必然联系的理由

我前面提到过，有些哲学家(最著名的要数布伦坦诺)赋予心灵以一种"意向性"的力量，正是它使心灵能够去进行指称。显然，我不同意这种观点，因为它不解决问题。我这样做的根据何在呢？我是不是可能有点过于武断了呢？

这些哲学家并没有主张我们根本不用表征也能思考外部事物或属性。我上面给出的那个论证，即把视觉材料比作蚂蚁的"图像"(以及通过以下科学幻想故事进行的论证：颜料溅出造成一棵树的"图像"，这幅画像产生出与我们"对树的视觉意象"有同样性质的感觉材料、但并没有树的概念与之相随)，可以看成是表明：意象并不必定指称外物。如果有什么心理表征必定指称(外部事物)的话，它们也必定具有概念的本性，而不具有意象的本性。但什么是概念呢？

当我们进行内省时，我们觉察到的并不是在我们心里流逝着

① 如果钵中之脑与"未来的"树有因果联系，也许现在它们可以用"在将来某个时间我将指称为'树'的某物"这一表述来指称树。但是，我们要设想一个事例，其中钵中之脑从未到过钵外，因此，从未与树之类的有过因果联系。

的"概念"本身。随时随地打住思绪，我们把握住的只是语词、意象、感觉、情感等等。当我把自已的思想大声说出时，我并没有作两次思考。我和你一样听到我的语词。当然，我在说出我所相信的语词和我并不相信的语词时，感受是不同的(但有时，当我紧张不安时，或者面对一群不友好的听众时，就会觉得我是在说谎，尽管我知道我说的是真话)；我在说出我所理解的语词和我所不理解的语词时，感受也是不同的。但我可以毫无困难地想象，某人和我一样想到同样这些语词(也就是说在他心里说这些语词)并同样感觉到对这些语词的理解、断定等等，但过一会儿发现(或被一个催眠师唤醒)他根本没有理解刚才心里出现的东西，甚至不懂得这些语词所属的语言。我并不是说这很有可能，我只是说这并没有什么不可思议的。这段话的意思不是说概念就是语词(或意象、感觉，等等)，而是说，说某人有一个"概念"或一个"思想"，同说某人有任何心理"表征"、任何可内省的实体，是大不相同的。概念并不是以指称外部对象作为固有本性的心理表征，其决定性的根据就在于：它们根本不是心理表征。概念是以某种方式使用的记号；记号可以是公共的也可以是私人的，可以是心理实体也可以是物理实体，但即使记号是"心理的"和"私人的"，与它的使用无关的记号本身却并不是概念。记号本身并不具有指称的内在功能。

做一个非常简单的思想实验就可以看出这点。假如你像我一样，也不能辨别榆树和山毛榉。我们仍然说我话中的"榆树"的指称与其他任何人话中的"榆树"的指称一样，也就是说，指榆树，而且你和我的话中全部山毛榉的集合都是"山毛榉"的外延(也就是指用"山毛榉"一词对之适用的事物的集合)。在"榆树"所指称的和"山毛榉"所指称的之间的差别，是不是真的可以相信是由我们概念中的差别造成的呢？(我很惭愧地承认)我对于榆树的概念与我对于山毛榉的概念完全相同。(顺便说说，这表明指称的决定因素是社会的而不是个人的；你和我都不同于能够辨别榆树与山毛榉

的专家。)如果有人斗胆企图提出,在我话中"榆树"的指称和"山毛榉"的指称差别,可用我心理状态中的差别来说明,那么请他设想有一个在那里语词都对换使用的孪地球。孪地球与地球极为相似;事实上,除了"榆树"同"山毛榉"作了对换之外,那位读者可以设想孪地球同地球没有什么差别。假如在孪地球上我有一个同形人,他与我的每一个分子都同一(如果能对两根相同的领带使用"同一"一词的话)。假如你是一个二元论者,那么假定我的这位同形人和我具有同样的言语和思想、同样的感觉材料、同样的气质,等等。要认为他的心理状态同我的心理状态有一点差别都是荒谬的,但他的"榆树"一词表征山毛榉,而我的"榆树"一词表征榆树(同样,假如孪地球上的"水"是一种别的液体——比方说,XYZ,而不是H_2O——那么,孪地球上用的"水"字和地球上用的"水"字就表征不同的液体,如此等等)。和 17 世纪以来就一直存在着的一种理论相反,我们认为意义根本不在头脑之中。

我们看到,拥有概念不等于拥有意象(比方说,树的意象,甚至,在这个问题上,也包括句子或整段话语的意象,不管是视觉的,还是听觉的),因为,人们可以随意拥有任何意象系统,但不具有以适合情境的方式来使用句子的能力(考虑一下——前面讲过——,决定"情境的合适性"的不仅有语言因素,而且有非语言因素)。一个人可以具有他所愿意的一切意象,但即使许多树在他面前,当要他"指出一棵树"时,却仍然不知所措。他甚至可以具有关于他所应当做的事情的意象,但依然不知道他应当做什么。因为,如果不伴随着以特定方式行动的能力,意象就只是一幅图像,而依据一幅图像进行行动本身,则是一种人们可有可无的能力。(此人可以形成一幅他本人正指着一棵树的图像,但这只是为了考虑某种逻辑上可能的事情;这幅图像所表征的是当某人发出了"请指出一棵树"的声音系列——对他来说这并无意义——之后,他对一棵树的指出。)他可能仍然不知道他应该指出一棵树,而且仍然不理解"指出

一棵树”。

我把使用某些句子的能力看成了是不是具有一个十足的概念的标准，但这个标准可以很容易地放宽。比方说我们可以容纳其要素不是自然语言中语词的符号系统，也可以容纳像意象这样的心理现象和其他各种内部事件。关键在于它们必须具有像自然语言中句子那样的复合性和互相结合的能力，等等。因为，虽然一个特定的表征——比方说，一道蓝色闪光——可以作为一个特定数学家对质数定理的整个证明的内部表达，一般也不大想这样说（这样说会是错误的），如果该数学家不能把他的“蓝色闪光”分解为若干不同步骤和逻辑联系的话。但是，不管我们把什么内部现象当作可能的思想表达，与上述论述完全相似的一系列论证将表明，构成理解的不是现象本身，而是思想者运用这些现象的能力，以及在合适的场合产生合适的现象的能力。

上述论证是维特根斯坦在《哲学研究》中所作论证的一个浓缩。如果这个论证是正确的，那么企图以所谓“现象学方法”去理解思想，完全是南辕北辙；因为，现象学家没有看到，他们所描绘的只是思想的内部表达，但对于该表达的理解——某人对自己思想的理解——并不是一个事件，而是一种能力。我们举的那个例子，即某人假装用日语思维（并且骗过了一位日本传心术士），已经表明用现象学方法研究理解问题的徒劳无益。因为，即使当且仅当某人真的理解，才有某种内省性质出现（事实上，根据内省看来并非如此），这种性质仍然只是同理解相关，那个愚弄了日本传心术士的人虽然也有这种性质，但仍有可能对日语一窍不通。

相反，请考虑这样一个完全可能存在的人，他并不具有任何“内心独白”。这人说得一口纯正英语，如果就给定论题请他发表意见，他会详细表达自己的观点。但是他从来都是要通过大声说话（用语词、意象等等）进行思考的；而且，除了听到他自己说话的声音，除了具有来自其周围的感觉印象再加上一种笼统的“理解感”

之外，并没有任何东西"经过他的头脑"。(或许他有同自己交谈的习惯。)在打一封信或去买东西等等的时候，他并没有一种内部的"思绪"；但是他的行动是有理智有目的的，随便什么人走上前去问他"你在干什么?"他都会作出完全一致的回答。

这个人看来是完全可以想象的。没有人会仅仅因为他进行有意识思考时总是大声说话，便不肯说他是有意识的人，他不喜欢摇滚乐(如果他经常表现出对摇滚乐的强烈反感的话)，等等。

上述一切告诉我们：(1)构成理解的不是心理事件(包括意象或更"抽象的"心理事件或性质)的集合；(2)理解的必要条件也不是心理事件的集合。尤其是，不能把概念等同于任何精神对象。因为，假定心理对象指的是可内省的东西，我们可以看到，不管它是什么，它可以不出现于某个对该相关词确实理解(因而具有十足的概念)的人之中，而出现于一个根本没有那个概念的人之中。

再回过头来看我们对神秘的指称理论所作的批评(这个论题维特根斯坦也关心)。我们看到，一方面，那些我们能够内省地觉察到的"心理对象"——语词、意象、情感，等等——就像蚂蚁的图像一样(基于同样的理由)，并不具有指称的本性。另一方面，企图设定一些确实与其所指有必然联系、而且只有训练有素的现象学家才能发现的特殊的心理对象即"概念"，也在犯一个严重的逻辑错误；因为，概念(至少部分地)是能力而不是事件，主张存在着必然地指称外部事物的心理表征不仅仅是蹩脚的自然科学；它也是蹩脚的现象学和概念混乱。

第二章　有关指称的一个问题

钵中之脑的假说原来是逻辑上非连贯的，为什么人们对此感到惊讶呢？原因在于，我们总是以为，我们头脑中发生的事件必然决定了我们想表达的意思以及我们的语词所指。但是不难看出，这种看法是错的，像我、这、此处、此刻这样一些普通指示词就是对这种看法的一种简单的反例。当我想到“我上班迟到了”的时候，我的心理状态同亨利（如果你愿意，可以设想亨利和我是完全相同的孪生子）的心理状态完全相同，然而，出现于我思想中的“我”字这个标志指称着我，出现于亨利思想中的“我”字这个标志则指称着亨利。我在星期二想到“我上班迟到了”，与我在星期三想到“我上班迟到了”所处的心理状态可以是相同的①，但我所说的“迟到了”所指的时间在这两例中并不一样，自然物种术语的例子说明的也是这个意思，但更加微妙一些。

把上一章提到的例子说明白点。假定在孪地球上有讲英语的人（由于一种不可思议的巧合，他们恰好进化得同我们毫无二致，并且所讲的语言除了我将要提到的差别之外，同两百年前的英语

① 至少我可以在以下意义上处于“相同心理状态”，即包括在导致我想到那个想法的心理过程中的参数可以具有相同的值。固然，我的总的心理状态是不同的，因为在星期二我相信“此刻是星期二”，而在星期三我不这么相信；但是，如果一种理论断言语词的意义随个人的总的心理状态的变化而变化，这种理论将根本不会允许任何语词有相同意义，因而就等于把有关意义这个词的概念本身也抛弃了。此外，我们可以虚构一个孪地球的故事，其中我和我的同形人处于相同的总的心理状态，而“我”和“此刻”的指称仍然是不同的（孪地球上的历法与我们的不同）。

一模一样），我将假定这些人尚不知道道耳顿的化学或道耳顿以后的化学，所以，具体说来他们还没有像“H_2O”这样的概念。现在假定孪地球上的河流湖泊里充满着一种看上去像水但不是H_2O的液体，在这种情况下，孪地球上使用的“水”这个词指称的不是水，而是这种与水不同的液体（比方说，XYZ）。但是，在孪地球的说话者和地球（比方说1750年的）说话者的心理状态中，并没有与此有关的差别来说明指称上的差别。指称不同是因为材料不同[①]，与整个情况相分离的心理状态本身，并不确定指称。

然而，有些哲学家反对这个例子。这些哲学家提出，假如这样一颗行星真的被发现的话，就应该说“存在着两种水”，而不应该说我们所说的“水”这个语词不指称那种孪地球液体。在这些批评者看来，如果我们真的发现河流湖泊中充满着一种与H_2O不同、看上去像水一样的液体，那么我们就将证伪了所有水都是H_2O这个陈述。

修改这个例子以避开上述争论是不难的。首先，孪地球上的液体不必同水那么相像。假定这种液体实际上是一种20%的谷物酒精与80%的水的混合物，但是孪地球上人的机体的化学结构使他们并不喝醉，甚至尝不出这种混合液体与水之间的差别。这样一种液体可以与水有许多差别，但是一个有代表性的说话者对这些差别却可能一无所知，从而与地球上1750年的有代表性的说话者处于完全相同的心理状态。当然，对我们来说孪地球的“水”与地球的水味道不同；但是对他们来说，味道并无不同。这种“水”煮沸时呈现出不同性质；但是，难道一个说英语的人为了把一种相当标准的概念内容与“水”这个词相联系，就一定注意过水到底在什么时候沸腾，以及到底发生了什么情况吗？

---

① 若要了解这个问题的详细讨论，请参见“‘意义’的意义”，收于拙著《心灵、语言和实在》（《哲学论文集》，第1卷），剑桥大学出版社，1975年。

有人可能会提出异议说，很可能在孪地球上有一些专家确实知道关于"水"的情况（比如，它是两种液体的混合物），而有关水的此类情况我们并不知道（在1750年不相信水有这种情况，因为它们不适用于水），因此孪地球上讲英语的人的集体的心理状态同（1750年的）地球上讲英语的人的集体的心理状态是不同的。有人可能承认某人词项的指称不取决于他个人的心理状态，但又坚持说该语言的共同体所有成员的整个心理状态决定了该词项的指称。

这里有一个困难，即在1750年地球或孪地球上的人还不一定拥有那么多的化学知识。如果地球上化学建立起来之前该词项具有它在今天具有的（日常用法中）同样的意义与指称，如果孪地球上相应知识建立起来之前该词项具有它今天具有的相同的意义和指称，那么，我们就可以追溯到这个较早的时候，那时这两个社会的集体心理状态在有关确定"水"的外延的一切方面都是相同的，并且争辩说，那时（正如事实上现在）外延就是不同的，因而集体的心理状态并不确定外延。我们可以由此断言化学建立起来了指称就改变了吗？我们可以由此断言一个用来指称两种水（尽管我们感到味道不同）的词项在化学建立起来以后就仅仅指称其中一种水了吗？

如果我们说，孪地球人的词项或我们的词项在他们或我们建立了化学（在能够蒸馏液体，分辨出掺水酒精是混合物这一点上来说）之后就改变了指称，那么我们就得说几乎每个科学发现都改变了我们词项的指称。根据这种观点，我们并没有发现（前科学意义上的）水是 H_2O；相反，我们规定了水是 H_2O。在我看来，这种说法显然是错误的。我们一直把水理解成具有用该词项来辨认称呼的那种周围物质的性质的东西；而这样意义上的水，我们发现就是 H_2O；孪地球上的人们则一直把"水"理解成为用该词项在他们的环境中辨认称呼的那种物质，而他们的专家发现此意义上的"水"

是两种液体的混合物。

当专家们作出诸如"水是H_2O"或"水是两种液体的混合物"之类的发现时，"水"(在地球和孪地球的两种语言中)的意义并没有改变，或它的日常意义和指称并没有改变(当然，这样的发现会导致若干较专门用法的出现)，并且地球上"水"的意义和指称并不包括酒、水混合物，如果我们同意上述看法的话，那么，我们就得说，专门的知识既不能用来解释地球和孪地球上"水"这个词在意义上的差别，也不能用来解释指称；因为我们还可以考虑另外一个孪地球，在那儿水是另一种混合物，但那里的专门知识却同前一个孪地球的专门知识没有区别(都相当匮乏)。要不，就像刚才所说的，我们可以很简单地设想在地球和孪地球上还不存在专家。即使这两种共同体的集体的心理状态相同，"水"这个语词仍可以指称不同的物质。人们头脑中所发生的东西并不确定他们词项的指称。用穆勒的一句话来说："实体自身"确定了该词项的外延。

一旦我们看到心理状态(不管是个体的还是集体的)并不能确定指称，我们就不该为钵中之脑未能指称外物(即使他们的心理状态同我们的一样)从而不能说出或想到他们是钵中之脑而感到惊讶。

内涵、外延和"观念世界"

既然我们词项的指称不是简单地由我们的心理状态所确定，为了考察它到底如何确定，不妨就便利用某些专门术语。在逻辑中，一词项所适用的事物集称为该词项的外延。这样，词项"猫"的外延便是猫的集。如果一个词项的意思不止一个，我们就权当该词带有无形下标(这样实际上就是两个词而不是一个词)，比方说，"兔子$_1$"——外延：兔子的集——"兔子$_2$"——外延：胆小鬼的集。(严格说来，自然语言中的词项的外延总是有点模糊的；但是为了

简单起见，我们权当这种模棱两可的情况已经以某种方式作了规定。）

像“我”这样一个在不同场合其所指不同的词，并没有一个外延，而只有一个外延函项：这个函项在每个用词情况中规定一个外延。在“我”这样的词中，外延函项相当简单；它仅仅是这样一个函项 f(x)，即对任何说话者 x 来说，函项 f(x)的值就是由这种 x 本身所构成的集。涉及用来描述用词情境（在这里，也就是说话者）的相关参项的自变元“x”，在语义学中被称为指数。指数用来表示次数，表示用直指方式指称的事物，并且表示在一个完整的语义操作情境中的其他若干特征（但我们将不讨论细节）。

在不同的可能情境或“可能世界”中，构成“猫”的外延的事物集也不同。在不存在猫的可能世界 M 中，“猫”的外延是一个空集。假定我的猫爱尔莎有过后代的话，那么“猫”的外延就至少有一个它在现实世界中所没有的分子。（对此我们可以这样表达：在爱尔莎是有后代的各个可能世界 M 中，“猫”的外延包括若干它在现实世界中所不包括的分子。）

我们曾用一个函项来表明，“我”这个词的外延依说话者的不同而不同，我们可以用完全相同的方式来表明一个词项的外延依可能世界 M 的不同而不同。假定有一个由叫做“可能世界”的抽象对象构成的集，用来代表各种事态或可能世界的历史，并把一个函项 f(M)同“猫”这个词项相联，在各个可能世界 M 中，这个函项 f(M)的值是世界 M 中的猫这些可能对象的集。这个函项，我将按照蒙塔古和卡尔纳普的办法，称之为“猫”这个词的内涵[①]。同样，“接触”这个二元谓词的内涵是函项 f(M)，它在任何可能世界 M 中的值，是在世界 M 中互相接触的可能对象的有序对的集；三元

① R·蒙塔古：《形式哲学》，耶鲁大学，1974 年。“内涵”的这种用法不是我在“‘意义’的意义”中讨论的那种传统用法。

谓词"x 位于 y 和 z 之间"的内涵是这样的函项，它在任何可能世界中的值，是具有 x 位于 y 和 z 之间的序列的三元组(x,y,z)的集合，如此等等。一个像"我"这样的词的外延在任何世界中都是依赖于语境的，它的内涵将是一个比较复杂的函项，其自变元不但包括可能世界，而且包括代表语境的指数。

内涵的作用是具体规定外延依赖于可能世界的方式。这样，它很全面地代表了我们所关心的东西，也就是同某词项相关的外延，因为它道出了在任何可能世界中那个外延会是什么。

不能把(这种含义的)"内涵"等同于意义，理由是任何两个在逻辑上等价的词项在每个可能世界中具有相同的外延，因而具有相同的内涵，但是一个理论若不能在具有相同意义的若干词项与只是在逻辑和数学上等价的若干词项之间作出区分的话，就不是一个完全的意义理论。"立方体"和"正六面体"是两个逻辑上等价的谓词，所以这两个词项的内涵相同，也就是说，它们的内涵都是在任何可能世界中以该世界中立方体的集合为值的函项；但是它们在意义上有差别，如果我们简单地把意义等同于这个函项，就会忽视这种差别。

我要强调这一点，即在这一理论中，要把可能世界、集和函项看作抽象的超心理实体，而不要把它们混同于这些实体的表征或描述。

弗雷格曾认为，一表达式的意义(即含义)是可以被心灵以某种方式"把握"的超心灵实体或概念。我们对内涵作了上述理解之后，弗雷格的这种理论对我们便毫无用处。首先，如上所述，未被内涵把握的意义中有一些独特的东西；因此，对一词项的理解不能仅仅在于把它同一个内涵相连。更重要的是，如果假定我们并没有"第六感觉"去直接领悟超心灵实体，或从事类似于领悟(或说是"顿悟")它们的活动，那么，"把握"一内涵或任何超心灵实体，就必须以某种方式借助于表征的中介。(这从内省来说似乎也很清楚，

至少对我来说是清楚的。)但是我们要考察的问题是表征怎么能够使我们去指称心灵之外的东西。假定"把握"这个观念为心灵之外的X,乃是用未经证明的假定取消了全部问题。

如果我说到某人"相信桌子上有一杯水",通常我是认为他有指称水的能力。但是,我们已经看到,能够指称水,必须把这个词直接或间接地同现实的水(H_2O)相联系:"约翰相信他面前有杯水"这个陈述不仅是关于约翰头脑中发生的事情的陈述,而且在某种程度上也是关于约翰的环境以及约翰同该环境的关系的陈述。假如约翰原来是一个孪地球人,那么当他说"桌上有一杯水"的时候,他相信的是,桌上有一杯实际上由水和谷制酒精构成的液体。

如果我们想谈论某人头脑中发生的事情,而不想对该思想所指称的实际事物的存在或本质作任何假定的话,胡塞尔提出的一种方法是很有帮助的,那就是括号法[①]。我们说"约翰相信桌上有一杯水",如果我们把归之于约翰的那种信念加上括号,那么我们归之于约翰的就仅仅是一个(在完全"未加括号的"普通意义上)相信桌上有一杯水的现实的人或可能的人的心理状态。这样,假如孪地球上的约翰尝不出水和掺水酒精的区别,他可能和地球上一个现实的或可能的讲英语的人说"桌上有一杯水"时处于相同的心理状态,尽管他称为水的东西可以制成一份地道的掺水白兰地。我们将说他具有加了括号的信念(相信桌上有一杯水),实际上,括号法舍去了通常表达信念的用语中的若干言外之意(所有指称外部世界或思想者心灵外的东西的言外之意)。

丹尼尔·丹尼特近来以一种同胡塞尔使用括号法的方式相联系的方式使用"观念世界"这个用语[②]。按他的理解,一个思想者加

---

① 胡塞尔:《观念:纯粹心理学通论》,艾伦和昂温出版社,1969年。(初版于1913年。)

② D·丹尼特:"信仰之彼岸",载《思想和对象》,安德鲁·伍德菲尔德编,牛津大学出版社。

了括号的观念的总和构成了对该思想者之观念世界的描述。这样，孪地球上的人和我们具有相同的观念世界(甚至相同的观念水)；只不过他们生活在一个不同的实在行星上(从而指称不同的现实物质如“水”)。如果你愿意，前一章讲的钵中之脑也可以具有同我们在最小的细节方面都毫无二致的观念世界；只不过他们的所有词项都根本没有外部世界的指称。传统的意义理论主张思想者的观念世界决定了他们词项的内涵(而这些内涵，加上某特定的可能世界 M 是一个现实世界的事实，便决定了这些词项的外延和所有句子的真值)。我们已经看到，这种传统的意义理论是错误的；而这就是目前文献中包含了“意义”的许多不同概念(如“内涵”和“观念世界”)而不是一个统一的概念的缘故。“意义”已经支离破碎。但我们面临着拼合碎片的任务。假如内涵与外延不直接取决于观念世界，它们是怎样确定的呢？

我们语言的诠释对于诠释的公认观点是怎样由我们确定的(如果不是由个体确定，便是由集体确定)?对这个问题的最常见的观点同这样两个观念相关：操作制约和理论制约。起初人们对操作制约的设想是极为幼稚的；我们只要规定(事实上是约定)某句子(比方说，“电流通过这导线”)是真的，当且仅当某检验结果被观察到(电量计指针偏转，或用现象主义的话来说，我具有看到电量计指针偏转的视觉印象)。现在，没有什么人再捍卫这种粗拙的操作主义了，因为人们发觉(1)理论同经验间的关联是或然性的，不能把它们当作充分的相互联系加以正确的形式化(即使导线中有电流通过，总会有一些几率不高的事件或背景条件可能造成电量计指针不偏转)；(2)即使这些或然性的关联也不是简单语义上的相互联系，而依赖于有可能修正的经验理论。根据朴素操作主义的说明，每发现一种检验一实体是否真金的新方法，“黄金”的意义和指称都要发生一次变化。(实际上，我们不应该说什么发现一种新的

检验黄金的方法。）在操作主义者看来，理论是以句子为单位接受检验的（单个句子被规定的操作意义会告诉你如何检验该理论）；根据比较晚近的操作主义的描述，正如蒯因所说的，理论"作为一个整体面对经验的检验"。

然而，我们可以放宽操作制约的概念以克服或避开所有这些反对意见。这样，可以用形如"一个可接受的诠释使得句子S在实验条件E满足时的大部分时间为真"（或另一种形式："使得句子S在E满足时的大部分时间为假"）的制约条件对将被认为可接受的诠释（将内涵指派给某语言的谓词）的类加以限制。这种制约仿效了那种认为语言中句子的真假同经验之间有或然性关系的观点。其次，可以认为当理论发展时，这种制约有可能修正。不像朴素操作主义那样把它们当作意义规定，而把它们看作对于可接受诠释的类的尝试性限制，可以和皮尔士（他在布里奇曼宣布"操作作义"之前50年就已经写下了这样的话）一样认为操作制约的理想集本身并不是我们仅仅加以规定的东西，而是我们在经验探究的过程中不停地逼近的东西。简言之，可以认为，合理的探究者如果尽可能好地进行观察、实验和推理的话，他们应当作出的就是操作的制约，这些制约是他们处于"思虑平衡"状态即选出对我们的词项所作的诠释时，会加以采纳的；我们在任一给定时间实际上所接受的制约，具有一种合理估计或近似值的地位。

这种观点与蒯因坚持认为理论-经验的关联与我们知识整体的任何其他部分一样可以加以修正的观点是相容的。这种观点也并不把这样的修正看作"意义变化"；这样的修正在澄清我们原来一直在谈论的东西方面，在澄清早先的理论加操作制约仅仅一鳞半爪地把握的东西方面，可以作出而且常常确实作出好得多的努力。

除了用操作制约（或者朝着皮尔士那里的操作制约的理想集逐步逐步的逼近）来限制可接受诠释的类之外，还可以有一些涉及

理论的形式性质的制约。例如，“一个可接受的诠释应能使得不同的结果果真总是具有不同的原因。”康德认为这样的一个“理论制约”构成了理性本身：我们是把决定论原则强加于世界，而不是去发现它。以这种形式出现的制约当然太强了：维持决定论的代价将是我们作为整体的知识系统变得过于复杂。但是可以放宽这种制约，就像前面放宽操作制约一样。（例如，可以要求凡在以下情况就坚持决定论，即理论中其他部分在复杂性方面的“代价”不是太大；以这种形式出现的制约看来是我们通常接受的。）

理论制约常常被表述成对于理论选取的制约而不是对于理论诠释的制约，但可以很方便地对它作重新解释，使之起到后面那种作用。这样，如果某作者把“保守主义”或“保护主义”表述为对理论选取的一种制约（“不要选取一种以放弃许多既往的信念为条件的理论，如果已经有一个——在其他方面同样简单的——理论能维持这些信念并符合观察的话”），那么，我们就可以对它加以重新表述，使之成为如下对于诠释的制约：“一个可接受的诠释应能使已被长期选取的语句被认为是真的语句，除非这种诠释要求在该诠释下为真的语句集所构成的理论变得过分复杂，或要求对操作制约作大幅度修正，我们方可说该诠释是不可接受的。”此外，人们普遍认为，如果我们对特定观察材料下可以被接受的假说不作先验的排列（叫做“简单性排列”或“似真性排列”）的话，就不可能有归纳逻辑（尽管在不同的实验或观察情境中这种排列本身可能有所不同）；“在一个可接受诠释下为真的语句集在简单性排列中不得低于任何具有相同观察或实验后果的其他语句集”这样一个制约，就可能相应于归纳逻辑中的以下制约：应该选取同观察相符的诸假说中最简单的（或“似真性”最高的）假说。

科学哲学的文献中已经提出了许多种理论制约，有一些制约，比如“简单性”，涉及被选取的语句集的性质，也有一些制约涉及研究者选择该语句集的历史。但我们不必一一加以考察。之所以把

注意力集中于这样的观点，即我们语言的可接受的诠释（也就是对语言词项所作的可接受的内涵赋予）取决于操作制约和理论制约，理由是很明显的：某人是否具有某种经验这一点是心灵能够作出判断的（尽管关于"经验"有一些哲学上的问题）。所以，如果一理论蕴涵或包含一个由某种操作制约而同一经验 E 相联系（不管是不是或然性的联系）的语句，思考者看一看他是否具有经验 E，就可以知道至少在这种情况下，该理论是否有效、在融贯符合方面有没有麻烦。因为我们用来检验该理论的制约也确定了词项的外延，所以该思考者对于此理论之"有效性"的估价同时也是对它的真理性的估价，因为该说话者对于这些制约的知识就是对于词项的内涵的知识，所以，就所提出的任何关于这个世界的理论而言，把握正确的语义，人们就会知道如果此理论是真的，我们的世界就该是什么样的。

进一步说，如果我们理想化地假定，思考者们彼此间具有经济学家们所谓的"完全信息"，那么每个思考者就都知道已选取的理论问题的形式结构，它所从属的研究纲领的过去历史，它信守或不信守的信念，等等。这样，每个思考者都能够知道理论制约有没有符合。（假如我们不愿理想化地假定完全信息，我们仍然可以说思考者的整体能知道这一点。）

简言之，如果这个公认观点是正确的，我们就可以对内涵与外延如何确定作出出色的说明（原则上如此——当然，细节过于复杂，在目前的方法论知识的阶段上要全部提供是不可能的）。然而，不幸的是，这公认观点却行不通！

为什么公认观点行不通

公认观点的困难在于，它试图用确定整个语句的成真条件的办法来确定个别词项的内涵与外延。正如我们看到的，其中包括这

样的意思，即操作制约和理论制约（即合理的探究者在探究的某种理想极限内会选取的那些制约）决定了语言中哪些语句为真。然而，即使这种观点成立，这样的制约也不能决定我们的词项指称什么东西。因为，在操作制约和理论制约的概念中没有什么东西直接具有这种作用。然而，间接地做，即撇开选择真语句集的制约，然后希望，通过确定整个语句的真值而确定出现在这些语句中的词项指称什么，这种办法也不会成功。

蒯因指出过这办法行不通[1]。我将以一种非常极端的形式推广以前的"不确定性"结论。我将证明，即使我们具有确定一语言中任何语句在任何可能世界中的真值的不管什么性质的制约，个别词项的指称仍然是不确定的。事实上，有可能对整个语言作截然不同的诠释，其中每一个都符合这样一个要求：每个语句在每个可能世界中的真值应当是经过明确规定的，一句话，不仅那个公认观点不成立，而且任何仅仅确定整个语句的真值的观点都不能确定指称，即使它为语言明确了在每个可能世界中的真值。

详细的证明技术性较强，我想还是在附录中提出更合适一些，下面我要提供的只是对这个证明的方法的说明，而不是详细的证明。

请看以下句子：

(1)一只猫在一张席上。（此处及往后"在……上"没有固定时态，也就是说，它的意思是"现在，过去和将来都在……上"。）

根据标准的诠释，在那些于某个时刻（过去，现在或将来）存在着至少一只猫或至少一张席的可能世界中，这个句子为真。而且，

---

[1] 参见他的"本体论的相对性"，载于论文集《本体论的相对性》，哥伦比亚大学，1969年。

“猫”指称猫，“席”指称席。我将表明，句(1)可作重新诠释，使得在现实世界中“猫”指称樱桃、“席”指称树而对句(1)在任何可能世界中的真值毫无影响。(“在……上”原来的诠释不变。)

这就是说，对句(1)将作出一种新的诠释，在这种诠释中，它的意思成了这样：

(a)一只猫*在一张席*上。

成为一只猫*(或者一张席*)的定义由这样三种情况给出：

(a)某猫在某席上，并且某樱桃在某树上。

(b)某猫在某席上，并且无樱桃在任何树上。

(c)以上两种情况都不出现。

下面是对成为猫*和成为树*的两种属性的定义：

“猫*”的定义

x 是猫*，当且仅当情况(a)成立且 x 是一樱桃；或者情况(b)成立且 x 是一猫；或者情况(c)成立且 x 是一樱桃。

“席*”的定义

x 是席*，当且仅当情况(a)成立且 x 是一树；或者情况(b)成立且 x 是一席；或者情况(c)成立且 x 是一夸克。

现在让我们来看，在归于情况(a)之下的可能世界中，“一只猫在一张席上”是真的，“一只猫*在一张席*上”也是真的(因为在这种世界里，一樱桃在一树上，并且一切樱桃都是猫*，并且一切树都是席*)。因为现实世界中某樱桃在某树上，所以现实世界属于这种世界，所以现实世界中“猫*”指称樱桃，而“席*”指称树。

在归于情况(b)的可能世界中，“一只猫在一张席上”是真的，

“一只猫*在一张席*上”也是真的（因为在归于情况（b）之下的世界中，“猫”和“猫*”是外延相同的词项，“席”和“席*”也是这样）。（注意，虽然在某些世界中——归于情况（b）之下的世界中——猫是猫*，但在现实世界中它们并不是猫*。）

在归于情况（c）的可能世界中，“一只猫在一张席上”是假的，“一只猫*在一张席*上”也是假的（因为樱桃不可能在夸克上）。

总起来说，我们看到，在每一种可能世界中，一只猫在一张席上，当且仅当一只猫*在一张席*上。这样，把我们刚才赋予“猫*”的内涵赋予“猫”，从而对“猫”这个词作重新诠释；同时把我们刚才赋予“席*”的内涵赋予“席”，从而对“席”这个词作重新诠释——这样做只会带来这样的后果，即使得“一只猫在一张席上”的意义等于“一只猫*在一张席*上”被规定具有的意义；这同每个可能世界中“一只猫在一张席上”被赋予真值的方式是完全相容的。

在附录中，我表明可以对整个语言的所有语句进行这种性质的更加复杂的重新诠释。这意味着，对于一语言的谓词，总是存在着无穷多的不同诠释，它在一切可能世界中把“正确的真值”赋予语句，而不管这些“正确的”真值是怎样选择出来的。蒯因在其《词与物》中论证了一个相似的结论；在蒯因所举的（适用于英语的）例子中，“那儿有一只兔子”被重新诠释成具有这种意义：“那儿有一块兔子切片”（其中“兔子切片”是整个四维时空的兔子的三组空间截面），或用另一种方式，具有这样的意义：“又存在着被具体化的兔子性”（后面这个重新诠释对语句的句法形式或至少它的逻辑句法也作了重新诠释）。蒯因提出了我们刚才提出的论点，即整个语句的真值条件不决定指称。由于“兔子切片”、“兔子性”以及“未分离的兔子各部位”都同兔子具有密切联系，所以人们可能从《词与物》中得到这样的印象，即所有未改变一语句之真值的重新诠释同标准诠释至少是密切联系的（就好像兔子的部位，兔子性同兔子的联系一样）。但本书附录中详细提出、本章稍作介绍的这个论证，却

表明"一只猫在一张席上"的真值条件甚至对"猫"指称樱桃的可能性也不排除。

"内在的"与"外在的"

当一个人面临把"猫"诠释成"猫*"和把"席"诠释成"席*"的非标准诠释时，第一个想法是认为它们向我们提供了一个微不足道的悖论，从而把它撇开不管。但是，真正的悖论决不是微不足道的；它们总是表明我们一直在用的思想方法出了毛病。第二种想法可能是抗议说"猫*"和"席*"是一些"古怪的"性质；我们的词项所对应的肯定不是这些"奇怪"的性质而是(如"是一只猫"或"是一张席"这样的)"可感"性质。人们指出可以"造一台机器"来"端详"一些物体从而"认出"它们是不是猫(人就是这样一台"机器")，以此来澄清为什么说猫*(要不就是猫性*或猫*性)是一种奇怪的性质，但是不可能制造一台将(在任何同我们的世界具有相似的规律和一般条件的世界)认出某物是不是一只猫*的机器。假如该机器(或一个人)看了看某物，并看到它既不是猫也不是樱桃，他们就可以认出这不是猫*；但是假如此物不是猫便是樱桃，那么该装置或该人就需要被告知"一只猫在一张席上"和"一颗樱桃在一棵树上"的真值，以确定它正端详或看着的是不是一只猫*，而这些真值仅仅通过考察为让它端详而向它呈现的对象，是不可能学到的。

不幸的是，人们可以对"看到"作重新诠释(比方说，诠释成看到*)，从而(3)约翰(或不管什么人)看到一只猫和(4)约翰看到*一只猫*这两个句子在每一可能世界中将具有相同的真值(用附录中提供的方法)。因此，某人只要看到一只猫，他就在看到*一只猫*；我们看到猫时具有的典型经验，就是我们看到猫*时的典型经验，如此等等。同样，我们可以对"端详"和"认出"作重新诠释，使得一台机器在端详一只猫时，它就是在端详*一只猫*，当它"认

出”某物时，它就是在认出*一只猫*。

用(诺齐克提出的)一个例子来说，假定我们中的一半人(或许是女人)用“猫”去表示“猫*”的意义，用“席”表示“席*”的意义，用“看”去表示“看*”的意义，用“认出”表示“认出*”的意义，等等。假如其余的一半(男人)用“猫”指示猫，用“席”指示席，用“看”指示看，等等。我们怎么知道情况是不是这样呢？[①](如果你问一个男人“猫”指称什么，他会回答“当然是猫”，一个女人也会这样回答，而不管“猫”指称什么。)

关键在于，事实上人们要能建造一台机器去端详物体并认出它们是不是猫、要能把猫同猫*区别开来的话，先得有把握说“端详”和“认出”指称着端详和认出，而知道这两个词的指称如何确定，并不比知道“猫”的指称如何确定来得容易。有人可能会说，当我看一看某物并且心想它是一只猫时，我的“心灵表征”、视觉意象或触觉意象、言语化思想的“猫”等等，所指称的是猫性、是各种其他的物理性质或生物性质(具有某种形状，具有某种颜色，属于某个物种)，而不是其他动物。或许这种说法是真的，但这只不过重复了指称是以这种方式而不是以另一种方式确定的这个观点。这是我们想要加以说明的东西，而不是我们想要寻找的说明。

“但是，”有人会抗议说，“前面给出的‘猫*’和‘席*’的定义指称的物体并不是上述对象(树上的樱桃和席上的猫)，而是表示具有这些性质的对象的外在性质。在现实世界中，凡樱桃都是猫*；但是如果樱桃不在树上，那么即使樱桃的性质完全如故，它也可能不是猫*。相反，某物是不是猫则只取决于它的内在性质。

---

① 一个女人可能会回答说，当她说“猫”时指称的是猫*的假定在逻辑上不连贯(因为在她的语言里，她叫作“猫”的就是猫)。这个回答带来的安慰不大；它并没有排除这样的可能性：她叫作猫的就是男人们叫作猫*的，反之亦然；这就是诺齐克要表达的意思。

这里涉及的区分，也就是内在性质与外在性质的区分，能不能使我们对"奇怪的"诠释加以刻划并加以排除呢？

这种主张的困难在于在"猫"和"席"与"猫*"和"席*"的关系中有一种对称性。因此，假定我们对"樱桃*"和"树*"作这样的定义，使得在归于情况（a）之下的可能世界中樱桃*是猫，而树*是席；在归于情况（b）之下的可能世界中樱桃*是樱桃，而树*是树；在归于情况（c）之下的可能世界中樱桃*是猫，而树是光子。这样，我们就可以按如下方式用带有＊号的词项来定义"猫"和"席"：

情况(a*)，某猫*在某席*上，且某樱桃*在某树*上。
情况(b*)，某猫*在某席*上，且无樱桃在任何树上。
情况(c*)，以上两种情况均不出现。

奇怪得很，以上情况恰好是经过重新描述的原来的那些情况(a)，(b)，(c)。于是我们进行定义：

"猫"的定义

x 是猫，如果情况(a*)成立且 x 是樱桃*；或者情况(b*)成立且 x 是猫*；或者情况(c*)成立且 x 是樱桃。（注意：在这三种情况下猫结果还是猫。）

"席"的定义

x 是席，当且仅当情况(a*)成立且 x 是树*；或者情况(b*)成立且 x 是席*；或者情况(c*)成立且 x 是夸克*。（假定对夸克*作了这样的定义，它使得在第三种情况(c*)中夸克*是席，那么在这三种情况下席结果还是席。）

结论是：从一种把"猫*"、"席*"等等当作原始性质的语言的

角度来看，指称着"外在性质"（它们的定义提及 x 以外的对象）的是"猫"和"席"；而相对于"常规"语言，也就是用"猫"和"席"来指称猫性和席性（您明白我指的是哪些性质，亲爱的读者！）的语言，指称着"外在"性质的则是"猫*"和"席*"。说得更明白些，是"内在的"还是"外在的"，要根据人们把哪些性质选作基本的性质而定：性质本身无所谓内在还是外在。

"生存"和进化

眼下流行着这样一种观点，认为进化过程本身以某种方式造成了我们语词和心理表征同外部事物之间的一种符合；人们常说，如果不存在这样一种符合，我们就无法生存下来，而这种符合便是指称的关系，至少它们原始形式是这样。

但是"符合"和"指称"同生存到底有什么相干？同样，真理同生存有什么相干？

在此观点出现分歧。有些哲学家认为，如果我们的信念（中足够多的部分）不是真的，我们就无法生存。另一些哲学家则声称，即使我们那些最完善的理论也不是真的，至少我们没有理由认为它们是真的。托马斯·库恩主张我们的信念只指称这些信念之内的对象（有点像"哈姆雷特"只指称一出戏中的一个人一样）；库恩指出，用来说明科学之成功的是尝试-错误法，而不是它的对象与实在事物之间的符合。巴斯·范·法拉森在其近著中论证说，一个成功的理论不必是真的，而只求"在观察上是充分的"，即正确地预言观察。他也把科学上的成功（或"观察上的充足性"）解释为运用尝试-错误法的结果。

如果这些哲学家是对的，那么用进化为相信客观的指称关系辩护的全部思想就失去了基础。在这样的工具主义观点看来，进化仅仅在一些词项（观察词项）和"感觉的恒久可能性"之间建立了一

种符合[①]。这样的符合并不是指称，除非我们愿意放弃这样的观点：外部事物（可观察事物）不是仅仅由感觉组成的构造。

但是，我相信其他哲学家（那些主张如果我们的信念中没有足够多的部分是真的我们就无法生存的哲学家）是对的。

我相信这一点的理由是：尝试-错误法并不说明为什么我们的理论是"观察上充分的"；说明这一点的办法只能是考察对尝试-错误法为何成功加以说明的环境与人之间相互作用的各种特征。（尝试-错误法毕竟并不是在一切事业中都获得成功的！）假设这种相互作用在我们的心灵中产生恰好导致成功预见的假理论，等于假设一系列完全不可解释的巧合。但是，我们的信念（趋近地）是真的这个事实怎样说明我们的生存呢？

我们有些信念同行为密切相关。要是我相信"如果按下该键，我将得到某种我认为非常珍贵的东西"这个句子（假定我按通常方式理解此句，或至少把它同"括号内的""想象中的"这个通常信念相联系），那么，我将伸出手去按那个键钮。形如"如果我做 x，我将得到______"（其中空格表示活动者所具有的目标）的信念，叫做指令性信念。如果我们的指令性信念错了很多，我将要从事的行动也会有很多失败；所以，我们（足够多的）指令性信念之真，对于生存来说是必不可少的。

至于我们的指令性信念，它们本身是从许多其他信念中得来的。这些其他信念包括关于外部事物的性质特征和因果力量的信念，也包括我们自己的性格和力量的信念。如果这些信念基本错误，居然还能导致真的经验预见、真的指令性信念，那倒不可思议了。所以，既然我们（足够多的）指令性信念是真的，而这个事实最好解释成是因为我们其他信念（构成我们"关于日常世界的理论的信念"）至少是近于真的，我们就有根据相信，我们关于日常世界的

---

① 或译"对应"。——译者

理论至少是近于真实的，并相信，如若不然，我们就无法生存下去。

我们设想一下，我们中有些人实际上指称的事物是由非标准诠释J(见附录)指派给我们所用的词项的。这个诠释与对于指称我们的概念世界、我们的感觉、我们的意志等等的词项的标准诠释是一致的。所以，根据J和"常规"的诠释，"看来我按键钮"在按"括号中的意义"理解(即理解成我有某种自愿地按下键钮的主观体验)时，并没有相同的真值条件，但具有相同的诠释。"看来我得到了预期的满足"亦然。

这样，假如根据非标准诠释J我们有足够多的指令信念是真的，那么我们就有把握得到成功，就有把握生存(因为如果我们不是活着的话，我们就达不到这些目标)和延续后代(因为如果他们不是活着的话，他们就达不到这些目标)。简言之，(足够多的)指令性信念的J类之真对于"进化上的成功来说与I类之真同样有利。事实上，它就是I类之真，因为每个语句(不仅是指令性信念)的真值条件根据I和J都是相同的，根据I和根据J，我们指令性信念不仅仅与同样的主观体验相联；它们具有的真值条件也是相同的。从"进化"的观点(或愚见)来看，必不可少的是，在任何把这些信念与相关的行为联系起来的诠释中，我都有足够多的信念为真。进化会使我们产生一种具有(某些)真的信念的倾向；但这仅仅意味着，进化能赋予我们一种其语句(或同语句类似的东西)具有某种真值条件(以及某种行动条件，或"语言外规则")的表征系统，而进化对以语句为中介或以概念为中介的生存的影响，正是通过它的这样一种倾向实现的。但是，整个语句的真值条件，正如刚才指出的，并不确定语句中各部分的指称(加上"语言外规则"也无济于事，因为这些规则是根据J而成立的)。由此可见，认为进化在进行指称的表达式与外部对象之间确定了一种唯一的对应[1]关系(或哪怕是

① 或译"符合"。——译者

一个被认为合理的狭窄的对应关系域),根本是一个错误。

纯粹意向和非纯粹意向

我们已经知道,在我们的词项与外部的事物之间,自然界并没有确定任何一种符合关系。自然界要我们在运用语词和思想符号时采取这样一种方法,使得我们有足够多的指令性信念是真的,从而使我们有足够多的行为有利于我们"广泛的遗传的适应性";但这样一来,大部分指称就是不确定的了,蒯因极力主张,指称就是这么一回事——不确定!在他看来,认为有什么非常确定的东西与我们语言中的词项对应着,这种观点只是一种幻觉。他指出:

> 仍然以我们标准的严格有序的符号(它带有一个包含若干经过诠释的谓词的词汇,还带有若干量化变项的大致确定的值域)为例。即使对谓词作了无数次重新诠释、对变项的值域作了无数次修改,该语言中真的语句仍然是真的。实际上,通过对谓词进行恰当的重新诠释,可以使同样大小的任何值域得以适用;如果值域是无限的,就可以使任何无限值域得以适用;这便是司冠伦-勒文海姆定理。在所有这样的变化中,真语句依然是真语句。
>
> 或许,我们首先关心的应该是语句的真值和它们的真值条件,而不是词项的指称。

在下章,我将详细讨论这里提出的选择方案,即放弃迄今为止一直是整个讨论之前提的那个观点:语词处于同(不依赖于讨论的)事物和事物集之间的某种一一对应关系之中。但是,看起来有一条简单得多的出路:为什么不干脆说确定我们词项的指称的就是我们的隐含的或直接的意向呢?

在上章刚开始时，我就反对这个观点，认为这个回答没有告诉我们什么东西，因为它带有（某种相关的）预设指称能力的意向。现在该对这简单的评论作些发挥了。

问题在于，“意向”和“心理状态”这两个概念有某种模糊性。如果一个心理状态的出现和消失只取决于说话者“内部”发生的过程，我们且称之为“纯粹的”心理状态。这样，我有无疼痛只取决于我“内部”发生的事件，但是我是否知道雪是白的，不仅仅取决于在我“之内”发生的某个事件（即相信或确信雪是白的），而且取决于雪是不是白的，因而是在我的身心“之外”的某个事件。因而，疼痛是纯粹心理状态，而知识则是非纯粹心理状态。知识中有（纯粹）心理状态的成分，但也有根本不是心理性的成分，这个成分符合这样的条件，即某人之所信如果不是真的，就不是知识。我如果不处于一个恰当的纯粹的心理状态中，我便不处于知道雪是白的这个“状态”之中；但对于知道雪是白的这一点来说，处于一个恰当的纯粹的心理状态中，决不是充分的；还得有世界与之配合才行。

那么信念呢？我们已经对括号内信念（“观念世界”）作了定义，根据这个定义，具有一种括号内信念[桌上有水]，或具有一种包括桌上有水在内的观念世界，便是一种纯粹的心理状态。但是，根据前面曾说过的，相信桌上有水（不加任何“括号”）预设了此人所用的“水”一词实际上指称着水，而这依赖于某些“范式”的实际性质，依赖于此人同这些范式的直接或间接的因果关系，等等。当我具有关于桌上有水的信念时，在孪地球上我的同形人具有同样的括号内信念，但不是同样的信念，因为他用的“水”一词指称的是掺水酒精，而不是水。简言之，相信桌上有水，这是一种非纯粹的心理状态。（钵中之脑不可能处于这种状态，尽管它们也可以处于与之对应的“括号内”状态。）

适用于信念的，也适用于意向。意向的纯粹心理状态——例如，某人观念世界中用“水”这个词项来指称水的意向——同确定

现实世界的指称，毫无关系。意向的非纯粹心理状态——例如，用“水”这个词项来指称实际的水的意向——预设了指称（实在的）水的能力。

有些哲学家主张，可以用我所谓“括号内信念”的状态再加上指称，来定义信念，因此，

> 约翰相信雪是白的＝约翰相信[雪是白的]
>
> （也就是说在约翰的观念世界中，雪是白的）
>
> 并且在约翰的思想中“雪”和“白”这两个字（或他用来表达此信念的任何语词）分别指称雪和白这个性质。

对于相信雪是白的这是个什么样的事件，上述说明并没有提供一个正确全面的分析。尽管如此，这个说明中包括的如下观点还是可以接受的、肯定是正确的：相信的活动预设了指称的能力。完全在同样的意义上，意向的活动也预设了指称的能力！意向并不是造成语词的指称功能的心理事件：（在通常所谓“非纯粹的”意义上的）意向把指称当作一个为整体所必需的部分而包括在内了。用（非纯粹的）意向来说明指称，会导致循环。意向、相信等等这些纯粹的心理状态怎么能（在恰当的因果系列中）构成或造成指称，正是这个问题使我们一直感到十分困惑。

困惑之源

初看之下，我们的语词和心理表征有所指称，这是再明显不过的事情了。当我心想或者说出“那猫刚出去”时，这想法通常是有关我们的家猫米蒂的；在我想或说的句子中，“猫”这个词通常指称着米蒂是其中一员的那个实体集。但是，我们刚才看到，这个“有关”或指称的关系的本质，是使人困惑的。

这种困惑的部分原因，可在实在世界和观念世界之间的区别中找到，也可以在与之相应的信念与括号内信念之间的区别、意向和括号内意向之间的区别中找到。在发现存在着语言的非意向的"可接受诠释"(这里我所说的可接受诠释仅仅是指满足适当的操作性和理论性制约条件的诠释)时，我们感到很惊奇、很疑惑，其部分原因在于，在说话者的"观念世界"中，是没有这种"不确定性"的。在我们观念世界中，猫和猫*是截然不同的(实际上，在我们的观念世界中猫*是樱桃)。"席上有一只猫"和"席*上有一只猫*"在逻辑上可能是等价的，但它们却包含了观念性所指大不一样的词项；因此，一个信念的实在世界所指与另一个信念的实在世界所指之间居然会有任何混淆，确实是很奇怪的。

但是，如果猫的数目与樱桃的数目恰好一样，那么，根据模型论定理(如蒯因在前面所引的话中所指出的)，有一个对整个语言的重新诠释，它使得即使在置换"猫"和"樱桃"的外延时，所有语句也都不改变其原有真值。利用刚才提到的技术，可以把这种重新诠释构造得能保留全部操作性的或理论性的制约条件(利用我们用"猫/猫*"这个例子所阐明的技术，可以对这种重新诠释加以推广，从而能提供规定每个可能世界的外延，而不仅是现实世界中的若干外延的"意向"或函项)。这同前面关于我们的"观念世界"或主观信念系统所作的陈述并不矛盾。因为，在我们的信念系统或"观念世界"中，猫都不是樱桃这个事实意味着，在对该信念系统的每个可接受的诠释中(在赋予词项、意象和我们在思想中运用的其他表征以外部世界所指的活动中)，"猫"的所指和"樱桃"的所指必须是两个不交叉的集。但是，这两个集的不交叉，同以下引人注目的事实形成了对照，即在这个可接受诠释中"猫"的集可能是另一个(同样可接受的)诠释中"樱桃"的集。观念中的猫或观念中的樱桃截然不同这个事实只意味着，如果可接受诠释的数目正好是一个的话，实在的猫和实在的樱桃就是截然不同的两个东西。如果对整

个语言的可接受诠释不止一个(如果只根据操作性的和理论性的制约条件选出可接受的诠释,情况就是如此),当考虑所有可能接受诠释的总和时,在各个可接受诠释中指称不交叉集的两个词项,就可能具有相同的潜在的所指,从观念中的猫和观念中的樱桃之间毫无共同之处这个事实,并不能得出结论说存在着猫自身和樱桃自身这两个确定的不交叉集。

上述事实之所以令人困惑,是因为操作性的以及理论性的制约条件,是允许由现实的经验情境来确定某人的表征系统的唯一可接受诠释(或若干可接受诠释)的唯一自然方式。这样的制约条件能够在某种程度上确定某人语言中哪些语句是真的,哪些语句是假的;然而,真值条件与指称之间的脱节现象却依旧存在。

蒯因——正如我们曾谈到的——对这种脱节现象倒愿意容忍,并干脆承认指称就是不确定的。一位年轻哲学家即哈雷·费尔德[①]最近提出了另一种观点。在费尔德看来,指称是语词或心理表征和对象或对象集之间的一种"物理主义关系",即一种复杂的因果关系。费尔德主张,这种物理主义关系是什么,要由经验科学来发现。

但这个观点也有问题。假定可能有一个对指称的自然主义或物理主义的定义,就如费尔德极力主张的那样。假定:

(1)x 指称 y,当且仅当 x 与 y 具有关系 R

是真的,其中 R 是一种可用自然科学词汇而不用任何语义学概念来定义(即不能用"指称"或任何其他马上会造成循环定义的词项来定义)的关系。如果(1)是真的,并且在经验上有可证实性,那么

① 费尔德:"塔斯基的真理论",载于《哲学杂志》,第 69 卷。我在《意义和道德科学》(劳特利奇和基根·保罗出版社,1978 年)一书中讨论了费尔德的观点。

(1)这个语句，即使根据以下理论，即一旦操作性的以及理论性的制约条件规定了指称，指称就得到了确定(而且仅仅这样，指称才得到确定)，它本身也是真的。(1)这个语句可能是我们的"反省平衡"或世界的"理想极限"理论的一部分。

然而，如果指称是仅仅由操作性的和理论性的制约条件确定的，那么"x 与 y 具有 R"的指称本身就是不确定的，所以，知道(1)是真的也无济于事。对象语言的每一可接受模型将对应于在其中(1)能成立的元语言的一个模型；"x 与 y 具有 R"的诠释将确定"x 指称 y"的诠释。但是，这将仅仅是一个在每一可接受模型中的关系；用它来消减可接受模型的数目，根本是无济于事的。

当然，这决不是费尔德的本意。费尔德所主张的是，(a)在语词和事物或事物集之间有一个确定的唯一的关系；(b)这个关系在赋予(1)本身以真值时被用作指称关系。但是，仅仅说(1)这个语句，并不一定表达上述内容，正如我们刚才看到的；我们怎么才能学会表达费尔德打算表达的内容，这是一个难题。

我们先把上述难题放在一边，来看看认为按费尔德要我们采取的方式理解的(1)(即把它理解成是对语词与其所指之间确定的、唯一关系的描述)是真的这种观点，如果这样理解的(1)是真的，那么是什么使之为真的呢？假定在词与物之间有许多"对应关系"，甚至有许多满足我们的各种制约条件的对应关系，那么，是什么选出一种特定的对应关系 R 呢？并不是(1)在经验上的正确性；因为那是一个有关我们的操作性和理论性制约条件的问题。如同我们已经看到的，也不是我们的意向(相反，在确定我们的意向的所谓时，R 也参与了进来)。看来，R 所指称的事实，必定是一个具有形而上学的不可说明性的事实，是一种原始的、无法还原的、形而上学的真理。

这种原始的、不可还原的、形而上学的真理，如果存在的话，切不可混同于近来由索尔·克里普克引入的那种"具有形而上学的

必然性的"真理。

克里普克的观点同前面提到的关于自然物种词项(例如,动物、植物和矿物的物种的词项)的观点有密切的关系,其内容是,假定有以下事实:

(2)水是 H_2O

(也就是说,假定(2)在现实世界中为真),并且假定(克里普克这样指出)说话者们有这样的意向,即"水"这个词项应该只指称那种与现实的水的各种标准样本具有同样规律性行为和同样基本成分的事物(也就是说,说话者们即使在谈论假设事例或"可能世界"时也有这样的意向),那么,(2)在每个可能世界也一定为真;因为,描述一种并不是 H_2O、但与水具有某些共同之处的假设液体,仅仅是描述一种与水相似的假设液体,而不是描述一个在其中水不是 H_2O 的可能世界,水是 H_2O 这一点具有"形而上学的必然性"(在所有可能世界都真);但是用来说明这种"形而上学的必然性"的东西却平凡得很:化学和有关说话者的指称意向的事实。

如果存在着一种确定的物理主义关系 R(不管它能否在自然科学语言中用无穷多的词项加以定义),它就是指称(与我们是否或如何描述该关系无关),那么,这个事实本身不能是我们的指称意向的结果;相反,正如我们一再指出的,在确定我们的指称意向本身的所谓时,它也参与了进来。克里普克的"水是 H_2O"在所有可能世界中都为真的观点,即使现实世界中指称仅仅由操作性和理论性制约条件确定,也可以是对的;此观点预设了指称的概念,它并没有告诉我们指称是否确定,也没有告诉我们指称是什么。

在我看来,相信某种符合关系内在地就是指称(不是我们的操作性和理论性制约条件或我们的意向的结果,而是一种基本的形而上学的事实),无异于一种神秘的指称理论。根据这样一种观点,

指称本身成了洛克所说的"实体形式"(一个内在地具有某个名称的实体)。即使人们愿意思考这种不可说明的形而上学事实,这种形而上学观点所伴随的认识论问题看来也是不可克服的。因为,假定有一个不依赖于心灵的世界,有一些不依赖于话语的实体(我们正在讨论的观点所预设的便是这种主张),那么正如我们已经看到的,就存在着表征可能的或可选择的指称关系的许多不同的"符合关系"(如果宇宙中存在着无穷多事物的话,实际上有无穷多这种符合关系)。正如我们已经看到的,如果(1)本身是经验地可接受(即可在我们的操作性和理论性制约条件的既定前提下接受)的话,那么,即使要求(1)根据任何形而上学地选出的"真实的"指称关系的真理概念是真的,也不排除这些可选择的指称关系中的任何一个。但这样一来,就存在着具有"R是真实的(形而上学地选出的)指称关系"这种形式的无穷多彼此不同的、可能的"不可还原的形而上学真理"。假如持这种观点的人承认他的观点并非十全十美,承认指称是形而上学地选出但又不完全是确定的(即形而上学地选出的R可以允许有多种可接受诠释),那么,就算把操作性和理论性制约条件的观点当作是形而上学地正确的这种看法,也是在意料之中的。因为,指称就是如下关系:x在至少一个可接受的模型M中指称y,为什么这不能成为一个不可还原的形而上学事实呢?要知道,所有这些无穷多的形而上学理论,与同样一些真语句,同样的"世界理论",和同样的发现何者为真的最佳方法论,都是相容的!

第三章 两种哲学观点

我们一直在讨论的那些问题自然而然地导致了两种哲学观点（或如我在序言中所称呼的，两种哲学气质）。我将讨论的就是这两种观点，以及这两种观点在几乎每个哲学争论问题上所造成的结果。“钵中之脑”的问题如果不是因为它揭露这两种哲学视角之间差异的尖锐方式，就毫无意义，当然，作为一种逻辑悖论除外。

这两种哲学观点之一，是形而上学实在论的观点。根据这种观点，世界是由不依赖于心灵之对象的某种确定的总和构成的。对“世界的存在方式”，只有一个真实的、全面的描述。真理不外乎在语词或思想符号与外部事物和事物集之间的某种符合关系。这种观点我将称为“外在论”观点，因为它最推崇的是一种上帝的眼光。

我要捍卫的那种观点还没有一个明确的名称。在哲学史上，这种观点出现得较晚，即便在今天，人们还常常将它与其他一些根本不同的观点搅在一起。我将称这种观点为内在论观点，因为这种观点的特征在于，在它看来，构成世界的对象是什么这个问题，只有在某个理论或某种描述之内提出，才有意义。许多（尽管不是所有）“内在论”哲学家还进一步主张，对世界的“真的”理论或描述不止一个。在内在论者看来，“真理”是某种（理想化的）合理的可接受性——是我们的诸信念之间、我们的信念同我们的经验之间的某种理想的融贯（因为那些经验本身在我们信念系统中得到了表征）——而不是我们的信念同不依赖于心灵或不依赖于话语的“事

态”之间的符合。并不存在我们能知道或能有效地想象的上帝的眼光；存在着的只是现实的人的各种看法，这些现实的人思考着他们的理论或描述为之服务的各种利益和目的。（“真理的融贯论”；“非实在论”；“证实主义”；“多元论”；“实用主义”这些术语都已被用在内在论观点身上；但是这些术语在历史上有一些别的用法，因而都带有一些不可接受的涵义。）

内在论的哲学家摒弃“钵中之脑”这个假设。对我们来说，“营养钵世界中之脑”不过是一个故事，一种纯粹的语言构造，而决不是一个可能的世界。显而易见，认为在某个宇宙，在另一种实在中此故事可能确有其事的人，在一开始就持一种上帝的眼光。因为，这个故事是从谁的角度出发讲述的呢？显然，并不是从世界之中任何有感知能力的生物的角度出发的。它不是从在另一世界中与该世界相互作用的任何观察者的角度出发的；因为，根据定义，一个“世界”所包括的每一样东西都以某种方式与其中的事物发生相互作用。比方说，假如你是一位并非钵中之脑的观察者，在观察钵中的那些大脑时，这世界上的有感知能力的生物就不能说全是钵中之脑。因为，那种假定可能有一个世界、其中所有有感知的生物都是钵中之脑的人，从一开始就预设了一种上帝眼光的真理观，或说得更确切些，一种没有人的眼光的真理观——把真理看成是同观察者完全无关的。

但是，另一方面，对外在论哲学家来说，要摒弃我们全是钵中之脑的假设就不是那么容易了。因为，一理论之真不在于它契合于观察者（们）所看到的那个世界（在这个意义上“真”就不是一种关系了），而在于它符合这自在世界本身。而我向外在论哲学家提出的问题则是，假如他是一个钵中之脑的话，他就不可能合乎逻辑地具有真和指称所赖以成立的这种符合关系本身。所以，假如我们是钵中之脑，我们也不可能想到我们是钵中之脑，除非是在加括号的意义上[我们是钵中之脑]；而这种加括号思想不具备能使之为真

的指称条件。因此我们是钵中之脑这一点是决不可能的。

假如我们持有神秘的指称论，比方说，我们可以主张有某些玄妙的射线——且称之为“新生”射线[①]——将语词和思想符号与它们的所指连接起来，那样就不成问题了。钵中之脑能够想到“我是一个钵中之脑”这些话了，而在他这样想时，“钵”这个词（借助于新生射线）符合于实在的外在的钵，而“之中”这个词（借助于新生射线）则符合于实在的空间容纳关系。但这种观点显然是站不住脚的。当今的哲学家中没有哪一个会相信这种观点。正因为现代的实在论希望在持有一种真理符合论的同时不相信“新生射线”（也不相信自我确认的对象[②]——内在地同这个而非那个语词或思想符号相符合的对象），钵中之脑这个事例才使他更为大惑不解。

正如我们已经看到的，要解决的问题是：在我们之外有这些对象，在我们之内则有正在进行思考/计算的心灵/大脑。思想者的符号（或他的心灵/大脑的符号）是怎样同外部对象或外部对象集发生一种一一对应的符合关系的呢？

在当今外在论者中流行的回答是，虽然记号并非必然地同这一组对象相符而不同另一组对象相符，然而，记号同外部世界之间的语境联系（尤其是因果联系）却能够使人阐明指称的本质。这个观点不能成立。例如，我关于电子的信念的主要原因可能是各种教科书。然而，我所使用的“电子”一词尽管在上述意义上同教科书有密切联系，却并不指称教科书。有些对象虽然是我的包含某个记号的信念的主要原因，但可能并不是该记号的所指。

外在论者于是会这样回答：造成“电子”这词同教科书之间联系的并不是一种“恰当类型”的因果之链。（但是，我们如不是已经

① “新生射线”是泽马赫向我提议的。

② “自我确认的对象”一词出自大卫·韦金斯的《实体和同一》（布莱克韦尔，1980年）一书。

能够指称，怎么可能想到去确定哪一种因果之链为"恰当类型"呢？）

对于像我这样的内在论者来说，情形截然不同。内在论也不主张记号会内在地符合对象，而不管这些记号如何运用、被谁运用。但是，被某一特定的记号使用者共同体按特定方式实际运用的一个符号，是能够在这些使用者的概念框架之内符合特定对象的。"对象"并不独立于概念框架而存在。我们在引入这个或那个描述框架时，就把世界划分为诸多对象。既然对象和记号同样是内在于此描述框架的，就有可能说什么和什么相符了。

说实在的，如果在某个语词所属的语言之内，用这个语词本身来说明它指称什么，意义并不大。"兔子"指称什么？噢，当然是指称兔子！那么"外星人"指称什么呢？指称外星人（如果真有外星人的话）。

外在论者自然赞成说"兔子"的外延是兔子的集，"外星人"的外延是外星人的集。但是他并不认为这样一些陈述在告诉我们指称是什么。对他来说，找出指称是什么，亦即，词与物之间的"符合"的本性是什么，则是一个紧迫的问题。（怎么个紧迫法，我们在上一章已经看到。）对我来说，关于一个概念系统中指称是什么，除了上述重言式命题之外，没有什么可说的。那种认为必须有一种因果关系的观点也被如下事实驳倒了：不管我们是否同外星人有过因果关系，"外星人"理所当然地指称外星人！

但是，外在论哲学家会答道，即使我们从未（就我们所知）同任何外星人有过交往，我们也能够指称外星人。因为，我们同地球人有过交往，我们曾经验过数例"来自别的行星"这种关系和数例"智慧生物"这种属性。我们能把外星人定义为来自地球以外的行星的智慧生物。而且，我们还可以把"来自别的行星"分析为"来自别的地方"和"行星"这两个词项（它们还可以进一步分析）。这样，外在论者便不再要求我们同我们能够指称的任何东西之间存在某种

"实在"联系(例如因果联系),而只要求基本词项指称我们与其有某种实在联系的那些物种(以及关系)。他说,通过运用复合命题中的基本词项,我们可以构造一些指称我们并未实际接触、甚至可能并不存在的物种(例如外星人)的摹状[①]式。

事实上,在像"马"或"兔子"这样简单的词项那里,我们就已经可以看到,外延包括许多我们与之并无因果关系的事物(例如,将来的马和兔子,或从未同任何人类有过接触的马和兔子)。当我们使用"马"这个词时,我们不仅指称我们与之有实际联系的马,而且指称属于同类的所有其他物体。

然而,在这里我们必须指出,如果离开了告诉我们什么属性被视作相同、什么属性不被视作相同的概念系统,"属于同类"这个说法是毫无意义的。在某些方面,任何东西都和任何其他东西"属于同类"。说什么我们指称某些事物是因为它们通过"某种恰当的因果之链"与我们相联系;说什么我们指称另一些事物是因为它们同我们通过某种恰当的因果之链与之相联系的事物是"属于同类"的;说什么我们指称上述事物之外的事物是借助于"摹状式";所有这个复杂的故事与其说是假的,不如说是无意义的。我未曾接触过的马之所以与我曾经接触过的马"属于同类",是因为不仅后者、而且前者,都是马。形而上学的实在论所表述的这个问题,又一次使人觉得似乎从一开始就存在着所有这些对象本身,然后我用某种套马索套住其中一些(即我通过"某种恰当的因果之链"与之有实在联系的马),接着我得设法用我的词("马")不仅去覆盖那些已经套住的马,而且去覆盖因为在空间、时间等方面离得太远而无法套住的马。而对于这个假问题(依我所见是假问题)的"答案"——形而上学实在论的答案——则是说,词不仅自动地覆盖我已经套住的对象,而且自动地覆盖属于同类的对象——即属于同一物种本

① "摹状"(description)在其他地方被称为"描绘"或"描述"。——译者

身。但这也就等于宣称，世界上存在着一些自我确认之物，因为它想要说的就是，把事物分为各个种类的是世界，而不是进行思维的人。

在某种意义上我可以说，世界确实包含一些"自我确认之物"——但并不是在外在论者所主张的意义上。在我看来，如果"对象"本身既是被发现的，又是被创造的，既是经验中的客观因素、不依赖于我们意志的因素，又是我们在概念上的发明，那么，对象当然内在地归于某些标签之下；因为这些标签是我们用来构造一幅世界（其中包括的首先是那些对象）图景的工具。但是，这种"自我确认之物"并不是不依赖于心灵的；外在论者企图把世界设想成是由既不依赖于心灵，同时又是自我确认的对象所构成的。这是办不到的。

内在论和相对主义

内在论不是一种说"什么都行"的圆滑的相对主义。有一种观点认为，讨论我们的概念同完全不受概念化过程污染的某物是否"相符"，是有意义的。另一种观点则认为，每一个概念系统因此同其他概念系统都恰好一般好。否定前一种观点是一回事，主张后一种观点又是另一回事。假如有些人真地相信那一套，假如这些人愚蠢之至，选择了对他们说他们会飞翔的概念系统，并依此行动，跳出窗外，如果他们侥幸保命的话，他们就会立刻发现后一种观点的弱点了。内在论并不否认知识有一些经验输入；知识并不是一种除了内在的融贯便一无制约的虚构故事；但是，内在论确实否认存在着任何其本身不在某种程度上为我们的概念、我们用于报告和描述它们的词汇所塑造的输入，否认存在着任何只允许有一种描述而同所有概念选择无关的输入。即使我们自己的感觉（它们作为知识的出发点对于一代又一代认识论家来说是相当重要的），我们对

它们的描述也受到大量概念选择的严重影响(在这方面,就像感觉本身一样)。我们的知识赖以建立的这些输入本身,总是受到概念污染的;但是受污染的输入总比没有输入强。如果我们只能拥有受污染的输入,那么我们只能拥有的东西仍然被证明是不可忽视的。

使得一个陈述,或整个陈述系统——一个理论或概念框架——具有合理的可接受性的,多半是它的融贯和契合。这种融贯包括理论性的信念或经验成分较少的信念互相之间的融贯,它们同经验成分较多的信念之间的融贯,以及经验信念同理论信念的融贯。我们对于融贯和可接受性的观点,根据我将要发挥的见解,同我们的心理难分难离。它们依赖于我们的生物组织和我们的文化;它们决不是"无价值"的。但它们确是我们的观点,对某些真实之物的观点。它们规定了一种客观性,一种对我们而言的客观性,尽管它并不是上帝的眼光那种形而上的客观性。从人的角度讲,客观性和合理性是为我们所拥有的;它们总比没有客观性和合理性要好。

拒绝那种认为有一种融贯的"外部的"视角的观点,那种认为理论"本身"就是真的、而与任何可能的观察者无关的观点,并不是要把真等同于合理的可接受性。真不能直接成为合理的可接受性,有一个根本的理由;真应当是一陈述不可或缺的一种性质,而被判明正当则是一陈述可以失去的性质。"地球是平的"这个陈述在3000年以前很可能是可以合理地接受的;但在今天却不能被合理地接受了。但是,如果说"地球是平的"这个陈述在3000年以前是真的,那就不对了,因为那会意味着地球已经改变了形状。事实上,合理的可接受性既是有时间性的,又是相对于某个人的。此外,合理的可接受性有一个程度问题;有时候也说真有程度问题(例如,有时我们说,"地球呈球形"大致是真的);但这儿的"程度"是指该陈述的精确性,而不是指它的可接受性或被判明正当的程度。

在我看来,以上所说并未表明外在论观点还可以被看作是对的,而只表明真是合理的可接受性的理想化。在谈论时我们假定存

在着认识论上的理想条件之类东西，一陈述如果根据此类条件被证明得以成立，我们就称之为“真”的。当然，“认识论上的理想条件”就像“无摩擦平面”一样：我们不可能真地得到认识论上的理想条件，甚至也不能绝对保证我们已经充分接近了这种条件。但是无摩擦平面也不可能真地得到，然而讨论无摩擦平面还是有“兑现价值”的，因为我们对它们的接近，可以达到一个相当高的程度。

用根据理想条件而判明正当来说明真的含义，或许会显得是用一个含糊概念来说明一个明确概念。但是当我们撇开诸如“雪是白的”这一类陈腐例子，“真”就不是那么明确了。不管怎么样，我要给出的不是正式的真的定义，而只是对此概念的一个非正式的澄清。

除了无摩擦平面的比喻之外，真理的理想化理论有两个主要观点：(1)真理不依赖于此时此地的被判明正当，但并非不依赖于时时、处处的被判明正当。宣称一陈述为真即是宣称它能被判明正当。(2)人们心目中的真理将是稳定的或“趋同的”；如果一陈述和它的否定都可以被判明正当，即使条件的理想程度尽如人意，以为该陈述具有一真值，也仍然毫无意义。

“相 似”论

主张真即符合的理论当然是最自然不过的理论。康德以前，要找出任何一位不持有真理符合论的哲学家，是不大可能的。

迈克尔·达米特[①] 近来区别了非实在论(即我所谓的“内在论”)观点和还原论观点，以便指出还原论者可能成为形而上学实

---

① 达米特的观点见《真理和其他谜》(哈佛，1980年)一书中“什么是意义理论I，II”一文。他随后(终于)作的威廉·詹姆士演讲(1976年作于哈佛)对这些观点作了较详细的发挥。

在论者，即真理的符合论的信奉者。就一类论断（关于心理事件的论断）而言，还原论主张该类中的论断是否真，取决于此类之外的事实。例如，根据某一种还原论的观点，使得关于心理事件的论断"为真"的，是关于行为的事实。又如，贝克莱主教认为一切"真实的存在"都是心灵及其感觉，也是还原论观点，因为它主张关于桌子、椅子和其他普通"物质对象"的语句之真实际上取决于关于感觉的事实。

如果一个观点就某一种论断而言是还原论的，但只对那类进行还原的语句坚持真理符合论，则归根结底仍然是形而上学实在论的观点。一种真正非实在论的观点，无论在什么情况下都是非实在论的。

经常有人错误地认为，还原论哲学家就是非实在论者，达米特当然没犯这个错误。还原论哲学家与其他哲学家的分歧之处在于实际存在的是什么，而不在于真理观。只要我们避免了这个错误，就会发觉我刚才所说的观点，即在康德之前不可能找到一个哲学家不是形而上学实在论者（至少在把什么当作基本的不可还原的论断这一问题上），是更为站得住脚的。

真理的符合论的最古老形式，存在了大约 2000 年的形式，便是古代和中世纪哲学家们归诸亚里士多德的那个理论。我不能担保亚里士多德真地持有此说，但他所用的语言给人留下了这种印象。我将称之为"指称的相似论"；因为这个理论主张，我们心灵中的表征与它们所指称的外部对象之间的关系，实际上是一种相似关系。

这种理论，像现代的诸多理论一样，也运用了"表征"这个观念。这种表征，这种心灵对外物的意象，亚里士多德称为"幻象"。在亚里士多德看来，幻象与外物的关系（幻象借此向心灵表征外部对象）是这样的，即该幻象与该外部对象具有同样的形式。既然幻象和外部对象是相似的（具有相同形式），拥有此幻象的心灵也能直

接拥有那个外部对象的形式本身。[1]

亚里士多德本人是这样说的：幻象并不与对象共同拥有像"红"这样的性质（亦即，我们心灵中的红和对象的红实际上不是同一种性质），不可能共同拥有像这种可以被一种感官所知觉的性质；但幻象可以和对象共同拥有像长度或形状这种能够被不止一种感官所知觉的性质（它们与"单一可感者"相反，是一些"共同可感者"）。

在 17 世纪，这种相似论的适用范围开始受到限制，虽然亚里士多德已经对它作了不少限制。洛克和笛卡儿认为，对于像颜色、质地这样的"第二"性质，要说心理意象的性质和物理对象的性质实际上是同一回事，是荒唐的。洛克是一个微粒论者，也就是说，是一个主张物质的原子论的人；和近代物理学家一样，他感到在我对一块红布的表征中感性地呈现的红色所对应的东西，并不是这块布的一种简单的性质，而是一种非常复杂的禀赋或"能力"，即产生这种特殊的感觉（用心理物理学家的话来说，表现为"主观的红色"的感觉）的能力。这种能力又需要用使之能有选择地吸收和反射不同波长的光线的布的微观结构来说明，但这种说明在洛克的时代人们还不得而知。（这一类的说明牛顿已经提出了。）如果我们说具有这种微观结构对一块布来说就意味着"是红色的"，那么，不管主观的红的性质是什么，在我具有一种主观的红的感觉时我的心灵（甚至我的大脑）中所发生的事件，显然并不包括我心中（或脑中）"是红的"任何东西。一物理对象的使之成为一种红色之物的性质，同一心理事件的使之成为一例主观的红色的性质，是截然不同的。一块红布和一块红的印象并非真的相似。它们并不具有同一种形式。

洛克的微粒论哲学使他将形状、运动、位置等视为基本的和不可还原的性质；对于这些性质，洛克愿意仍然运用指称的相似论。

---

① 见《论灵魂》，第 3 卷，第 7、8 章。

(诚然,有些研究洛克的学者现在对此意见不一;但洛克确实说在第一性质方面对象和观念之间有一种"相似",而红和热的观念与对象中的红或热之间,则没有相似。[1] 我这里所介绍的对洛克著作的读解方法,在他的同时代人和18世纪的读者中也很普遍。)

贝克莱的杰作

贝克莱发现指称的相似论有一个非常讨厌的结论:它意味着除了心理实体("灵魂"及其观念,即心灵及其感觉)便一无所有。贝克莱由之出发的前提——相似论——并不是他仅仅受自洛克(或加于洛克)的思想,而是在他以前就被公认,甚至在他身后又存在了一百年的指称理论;对此人们普遍地感到不以为然。但是我们刚才已经指出,这个理论确实有相当悠久的历史。

贝克莱的论证很简单,在第二性质方面哲学家们对相似论通常所作的反驳(即从感觉的相对性所作的反驳)是正确的,但它在第一性质方面同样行之有效。不同的感知者或同一个感知者在不同的场合对长度、形状、物体的运动的感知都是各各不同的。一张桌子同我对它的意象或你对它的意象是不是一样长,是一个荒谬的问题。如果该桌子3英尺长,而我对它的观察相当清晰,难道我就有一种3英尺长的心理意象么?这问题一提便知其毫无意义。心理意象并不具有物理长度,它们无法同巴黎的标准量尺进行比较。物理长度与主观长度之间,正像物理的红色与主观的红色之间一样,存在着重大区别。

贝克莱的结论换句话说就是:同某一感觉或意象相似的,只能是另一感觉或意象。假定这点,再假定(仍然是不成疑问的)指称的机制在于我们的"观念"(即我们的映象或幻象)与它们所表征的东

① 参见《人类理智论》,第2卷,第8章。

西之间的相似，马上就可以得出如下结论："观念"（心理映象）只能表征或指称另一意象或感觉。我们能思考、设想、指称的，只能是现象性对象。如果你无法思议某物，你就不能认为它存在着。我们对物理对象的谈论，除非被当作是关于我们感觉中的规则性的相当间接的谈论，否则就是完全莫明其妙的。

在贝克莱生前和身后，人们把他看作几乎是一个精神反常、离经叛道的人，一个即使才华横溢、但也居心不良的人。对他作如此评价的倾向，根源在于他那物质实际上并不存在（除非作为得自感觉的构造）的结论人们无法接受，而不在于他的前提有什么特别之处。但是从相似论出发居然能引出这样一个无法接受的结论，这一点引起了哲学中的一场危机。那些不愿追随贝克莱信奉主观唯心论的哲学家，不得不对指称作另一番说明。

康德论知识与真理

我想说的是，我们读康德时最好把他理解成第一次提出我所谓的"内在论的"或"内在实在论"的真理观的人，尽管康德本人从未道破这一点。

首先，康德显然认为贝克莱的主观唯心论是完全无法接受的（对这一点他说得颇为明白），而且认为因果实在论——即认为我们直接感知的只是感觉，对物质对象是通过某种颇成问题的推论过程而*推得的*观点——是同样无法接受的。康德说，认为在我写这页字时面前有一张桌子只是一个非常可疑的假说，这种观点是胡扯。

其次，我认为康德对贝克莱的论证过程看得很清楚：他看出这个论证以指称的相似论为依据，因而要拒绝贝克莱的论证就得拒绝这个理论。我在这里归诸康德的观点他并没有用这些话表述（实际上，把"指称"说成是心理记号同它们所代表的东西之间的关系

是很晚的事，虽然心理记号和它们所代表的东西之间的关系的问题古已有之）。但是我们将看到，康德实际上所说的内容具有的效果，恰恰是放弃指称的相似论。

请允许我提议用这样一种方法阅读康德的著作，它可能对我们有所助益，尽管这只是一种初步近似的正确诠释。试设想康德同意贝克莱的如下观点，即从知觉的相对性所进行的论证既适用于第二性质，也适用于所谓"第一"性质，但他对这个观点提出了与贝克莱不同的对策。贝克莱对策，刚才讲过，是取消第一性质和第二性质之间的区别，回过来求助于某种作为我们能加以指称的基本实体的东西，而这种东西正是洛克会称为"简单的"感觉性质的东西。我们刚才讲过，洛克本人对第二性质的处理，是说对这些物理对象的性质我们只能把它们设想为一种"能力"或能使对象按特定方式施加影响于我们（但其本性不详）的性质。说某物是红的，或热的，毛茸茸的，就是说它以如此这般的方式与我们相关，而不是说从上帝的眼光来看它如何如何。

我提议，按如下作为初级近似的方式阅读康德：指出洛克关于第二性质所说的话适用于所有性质——简单性质，第一性质，第二性质一视同仁（说实在的，对它们进行区别意义并不大）。[①]

如果一切性质都是第二性质，结论如何呢？结论是，我们关于一对象所说的任何东西都具有如下形式：它是怎样怎样，以至于能

---

① 康德在"序论"中正是以这种方式对他的观点作了概述。洛克之前很久，在他之后更甚，人们普遍地假定和以为外物的许多性质，可以说并不属于这些外物自身，而属于它们的显现，可以说在我们的表征之外，它们便没有严格意义上的存在，而这并不妨碍外物的实际存在。例如，热、色和味皆属此类。现在如我再往前跨一步，根据充分的理由把物体的其余性质，也就是像广延、位置和（一般地说）空间以及它所属的一切性质（不可入性，物质性和形状等等）这些所谓第一性质，也当作显现来对待，没有一个人能提出理由说这是不可接受的。不能把那种认为颜色不是对象自身的性质，而只是视觉的种种变化的人叫做唯心论者，也不能因为我发现构成对一物体之直观的更多性质，不，一切性质，都仅仅属于它的显现，便把我的观点称为唯心论观点。

以如此这般形式影响我们。我们关于任何对象所说的任何东西都没有描述对象的"自在"状态，都不能离开它对我们的影响，离开它对具有和我们一样的理性本性和生物构造的生物的影响。另一个结论是，我们不能假定我们关于一对象的观念和任何不依赖于心灵的、可对我们关于该对象的经验作最终说明的实在之间的相似性。我们关于对象的观念并不是不依赖于心灵之物的摹本。

康德正是以上述方式对问题进行描述的。他并不怀疑存在着某种不依赖于心灵的实在；对他来说，这实际上是理性的一种公设。他提到这种不依赖于心灵的实在的要素时，用过各种各样的术语：自在之物；本体对象或本体；作为总称，又叫做本体世界。但对这些本体之物我们不可能形成实在概念；即使本体世界的概念也是一种思想的界限(边界概念)，而不是一个明确的概念。今天，人们感到本体世界这个概念是康德思想中多余的形而上学成分。(但或许康德倒是对的：或许我们总要不由自主地认为我们的经验有某种不依赖于心灵的"根据"，虽然若想对它进行讨论便即刻导致无谓之语。)

与此同时，谈论普通的"经验"对象并不是谈论自在之物，而只是谈论为我之物。

真正微妙之处在于，康德认为所有上述观点不仅适用于外部对象，而且适用于感觉("内感官的对象")。这看来可能有些奇怪：一观念是否符合一感觉，这算什么问题呢?但康德看到了一些深藏着的东西。

假如我有一感觉 E。假定我通过像判定"E 是一种红的感觉"这样的方式来描述E。如果"红"只意味着"如同这个"，那么上述整个的论断就只意味着"E 如同这个"(同时指着E)，也就是说，"E 如同 E"——实际上什么判断也没有下。正如维特根斯坦指出的，到头来说的实际上是一串叽里咕噜。但是，如果"红"是一个真实的分类词，如果我说的是此感觉 E 与我在别的时间称为"红"的感觉

同属一类，那么我的判断就超出了直接给予的范围，超出"单纯所指"的范围，而包含了对别的感觉、对我当下所无之物的默认、对时间（在康德看来，时间并不是某种本体的东西，相反都是我们用以整理"为我之物"的一种形式）[①]的默认。要问我在不同时间具有的被归入红的感觉之类的感觉是否"真的"都（在本体上）彼此相似，是毫无意义的；如果它们看起来相似（例如，如果我记得以前的感觉同这个感觉相似，并预料我将归入此类的未来感觉也同这个感觉相似，如果这个感觉那时我还记得的话），那么，它们就是为我相似的。

康德一再（并以不同的术语）说道，内感官的对象并非先天地现实的（本体的），它们是先天地理想的（为我之物），它们同所谓"外部"对象一样，都不能被直接地认识。我称为"金块"的对象不能同本体对象直接比较以确定它们是否具有相同的本体性质，同样，我称为"红"的感觉，也不可能同本体对象直接比较以确定它们是否具有相同的本体性质。

之所以说"一切性质都是第二性质"只是初步近似于康德的观点，理由在于："一切性质都是第二性质"（即一切性质都是能力）意味着，说一张椅子是用松木做的，就是把一种能力（使我们觉得它是松木椅子的种种倾向）归于一个本体对象；说那把椅子是褐色的，是把另一种能力归于这同一个本体对象；如此等等。根据这样一种观点，康德所谓"表象"中的每个对象都有一个本体对象与之对应，也就是说，每个为我之物都有一个本体对象与之对应。但康德明确否认这一点。正是在这点上，他差点说他是不承认真理的符合论的。

---

① 这里我有意不顾年代顺序，用维特根斯坦《哲学研究》一书的一个例子来介绍康德的观点。但维特根斯坦的例子同康德也有密切关系：在康德之后不久著书立论并对他的学说一清二楚的黑格尔也有这样一个观点，即任何判断，即使是关于感觉印象的判断，如果真要成为一个判断的话，就必须超过所"给予"的东西。

诚然，康德并未说他不承认真理的符合论。相反他说，真理是“判断与其对象的符合”。但这是康德所说的“真理的名义定义”。在我看来，把它等同于形而上学实在论者对“真理的符合论”的理解，将是一个严重的错误，要断定康德是不是主张形而上学实在论者所说的“真理的符合论”，我们必须考察一下他对于他所谓的经验判断的对象，是不是持有一种实在论的观点。

根据康德的观点，任何关于外部对象或内部对象（物理之物或心理实体）的判断都这样说，本体世界作为一个整体如此这般，以至于一个理性生物（和我们具有相同的理性本性的生物），只要得到了一个与我们具有同样感官的生物（一个和我们具有相同的感性本性的生物）能得到的信息，就会构成这样一种描述。在这个意义上，此判断赋予了一种能力，但是此能力是赋予整个本体世界的；决不能以为，因为在我们的表象中有椅子、马匹和感觉，就相应地存在着本体的椅子、本体的马匹和本体的感觉。甚至为我之物和自在之物之间的一一对应也不存在。康德不仅放弃了关于在我们的观念和自在之物之间的相似性的任何概念，他甚至还放弃了关于抽象的同型性的任何概念。这就意味着，在他的哲学里，并没有真理的符合论。

那么什么是真判断呢？康德确实相信，我们是有客观知识的；我们有关于数学定律、几何定律、物理定律的知识，我们有关于个体对象——经验对象，为我之物——的种种陈述。运用“知识”这个术语和“客观”这个术语，便意味着断定仍然有一个真理的概念。然而，真理如果不是同事物的自在状态相符合的话，到底是什么呢？

我曾说过，从康德著作中能引出的唯一回答是这样的：一份知识（即一个“真陈述”）是这样一个陈述，只要和我们具有相同本性的生物实际上可能有的那种经验是充分的，一个理性生物就会对它加以接受。任何其他意义上的“真理”，我们都不可能达到，不可

能设想。真理是达到至善的契合。

经验论者的不同方案

在上述论证的范围内，一个哲学家若把真理的符合论和指称的相似论限于感觉和印象，以此避免对这两种观点的放弃，还是可能的。即使康德之后，许多哲学家依然相信我们之所以能够具有指称我们自己的感觉(以及别人的感觉，虽然这更容易引起争议)的观念，是因为利用了相似这个机制，相信从认识论的角度来看，最起码的指称便是相似。

为看出这观点为什么不能成立，且回顾一下，贝克莱论证的核心便是这样一个意思，即与某一个"观念"(感觉或印象)相像的只能是另一个"观念"，也就是说，在心理之物与物理之物之间，无相像可言。在贝克莱看来，我们的观念能同其他心理对象相像，但不能同"物质"相像。

在这里我们必须停一下，并指出上述观点有一个重要错误。实际上，任一物都在无数方面与他物相似。例如，我此时对打字机的感觉和我口袋里的那个硬币就它们的某些性质(那感觉出现于现在、那硬币此刻存在于口袋)都是我的过去行动的结果而言，是相似的。如果我刚才没有坐下来打字，我就不会有那个感觉；如果我以前没有把那硬币放在口袋里，它就不会在那儿。这感觉和这硬币都存在于 20 世纪。这感觉和这硬币都是用英语描绘的，等等，等等。如果你有足够的才智和时间，在任何两物之间可以找到无限的相似之处。

当然，在一个特定的情境中，"相似性"的意义可能更狭窄一些。但是在对正在争论的是何种相似性不加(直接地或含蓄地)具体说明的情况下，只是问"A 和 B 相似吗？"无异于问一个毫无内容的问题。

从这个简单的事实就可以看出，认为相似性是指称的隐秘机制的观点，一定会导致无穷倒退。举一个维特根斯坦提出的例子。假如某人想发明一种“私人语言”，一种指称他自己的直接给予的感觉的语言。他把注意力集中于一感觉 X，并引入一记号 E，这个记号他只想运用于那些在质上与 X 等同的实体。实际上，他认为应该把 E 仅仅用于那些与 X 相似的实体。

如果这就是他的全部意图，如果他对某物如能归入 E 类就必须同 X 在某个方面相似来作具体说明，那么，他的企图就是毫无内容的，这一点我们刚才已经看到了。因为，任一物都在某个方面与 X 相似。

但是，如果他对那特定的方面作了具体说明；如果他思考这样的想法，即一感觉是 E，当且仅当它在 R 方面同 X 相似；那么，既然他能够思考这个想法，他就已经能够指称他正设法为之引入词项 E 的那些感觉、指称那些感觉的相关性质了！但是，他怎么才能做到这一点呢？（如果我们回答说：“用这样的办法：把注意力集中于另外两个感觉 Z、W，并思考如下想法，即那两个感觉在 R 方面相似当且仅当它们同 Z、W 相似”，那么，我们便陷入了一种无穷倒退。）

指称的相似说所遇到的困难与我们先前提到的“恰当类型的因果之链”所遇到的困难如出一辙。如果我仅仅说“马”这个词指称具有如下性质的对象，即它的出现使我在某个场合作出“我面前有马”的陈述，那么就有一个困难，即这样的性质实在太多了。例如，假定 H-A（代表“马的显现”）是引起一位说一口流利英语的正常人作出“我面前有一匹马”这个反应的整个知觉情境。这样一来，当我说“我面前有一匹马”时（即使那时我正经历一种幻觉），H-A 的性质就发现了，但“马”并没有指称具有上述性质的情境，面是指称某些动物。一个具有属于某个特定物种这种性质的动物的发现，以及一种具有性质 H-A 的知觉情境的出现，都经过因果之链而同我

的陈述“我面前有一匹马”发生联系。事实上，在石器时代马的出现同我的陈述“我面前有一匹马”之间，就有着经过因果之链而发生的联系。正像指称如果仅仅是一个相似性的问题，就存在着许许多多相似性一样，指称如果仅仅是一个因果之链的问题，也存在着许许多多的因果之链。

然而，如果我说“马”这个词在某些场合指称具有同我作出“我面前有一匹马”这个陈述相关的性质的那些对象，是借助于一种合适的因果之链，那么，我面临的问题则是，如果我能够具体说明什么是这种合适类型的因果之链，我必须事先能够指称构成那种因果之链的物种和性质。但是，我怎么才能做到这一点呢？

结论并不是说不存在指称那些通过某些特殊种类的因果之链而同我们相联系的事物的词项，同样，结论也并不是说，不存在具有相似说所赋予的那种逻辑的词项。结论仅仅是，不管是相似性，还是因果联系，都不可能成为指称的唯一或根本的机制。

维特根斯坦论“遵守规则”

请考虑一下我刚才顺便提到的那个例子：一个人为了具体说明R这个方面（在这个方面感觉必须与X相似，如果它们要被正确地归入E类的话），打算通过口说或者心想两个东西在R方面相似的办法，来防止它们之间的相似方式等同于Z、W之间的相似方式。这个打算当然落空了，因此任何两物Z、W本身在不止一个方面相似（事实上，在无限多方面相似），企图用举出有很多例子的办法来具体说明一种相似性关系，无异于通过列举一自然数方程式的前1,000（或1,000,000）个值的办法来具体说明这个方程式：永远存在着无穷多的这样的方程式，它们符合关于任何有限的数值集的任何给定表格，但是不符合表上没有列出的数值。

这与维特根斯坦在《哲学研究》中提出的另一个观点有关，这

个观点我在第一章的末尾提到过。关于一个概念，不管我能够唤起的内省记号或“表征”是什么，它们都无法具体说明或构成该概念的内容。维特根斯坦曾在一个著名的小节中提出这个观点，这小节讨论的问题是“遵守规则”，比方说，“加一”这个规则。即使在两个可能世界中的两种生物（我尽可能以非维特根斯坦的术语陈述这个论证！）在“加一”这个文字表达的公式方面具有相同的心理记号，他们的实践仍然可能是不一的；而确定诠释的正是实践：我们前面看到，记号并不诠释自身。即使某人对“A 是 B 的后继数”这个关系（即 A=B+1）的描述和我们一样，并在大量但有限的实例（例如 2 是 1 的后继数，3 是 2 的后继数，……，999,978 是 999,977 的后继数）上与我们意见一致，他对“后继数”仍可能有另一种只有在某些未来的实例中才暴露出来的诠释。（即使他的“理论”——亦即，他关于“某某的后继数”所说的话——和我们一致，他仍然可能对整个理论作另一种诠释，正如司寇伦-勒文海姆定理所表明的。）

这不仅同语言哲学，而且同数学哲学，有直接关系。首先，有一个有限论的问题：人类实践，不管是现实的还是可能的，都只能在有限范围内扩展。即使我们说我们能够“连续不断地计数”，这也是不可能的。如果我们的实践有若干可能的不同范围，那么即使对自然数也有若干可能的不同诠释——我们的实践，或我们的心理表征等等，并不选取一个独一无二的自然数列的“标准模型”。我们之所以常常以为它们会作这样的选择，是因为我们很容易从“我们能够连续计数”转为“一台理想机器能连续计数”（或，“一个理想的心灵能够连续计数”）；但是，谈论理想机器（或心灵）与谈论实际的机器和人，是根本不同的。谈论一台理想机器能做什么，是在数学之内进行的，它并不能确定对于数学的诠释。

同样，维特根斯坦认为，谈论“相似性”、“同样的感觉”或“同样的经验”，是在心理学理论之内进行的；它并不能确定对于心理学

理论的诠释。而后者，即对于心理学理论和术语的诠释，是由我们实际的实践、实际的对错标准确定的。

纳尔森·古德曼在《世界的构成方式》[①] 中提出了一个与此密切相关的论点：企图对知觉事实在独立于我们用来把握它们的概念，独立于我们对它们的描述，独立于我们认为正当的描述的时候"实际是什么"得到一个概念，是徒劳无益的。于是，在讨论了心理学家苛勒的如下发现即相当数量的工程师和物理学家对于由不同位置上相继发出的光亮所造成的表象运动根本无法看出之后，古德曼评论道(第 92 页)：

> 然而，如果一个观察者报告说他看到两点分离的亮光(即使其空间间隔和时间间隔之短使得大多数观察者只看到一个运动光点)，他的意思或许是他看到这两点亮光就像我们看着一张椅子时说看到了一团分子一样，或者像我们斜视一张圆桌时却说看到了一个圆桌面一样。既然一观察者能够区别表象运动和真实运动，他也可以把运动的显现当作存在着两点亮光的记号，就像我们把桌面的椭圆形显现当作桌子是圆的记号一样。在这两种场合下，记号都是、或都变得如此显而易见，以至于我们可以通过它们而对物理事件或对象一目了然。当那个观察者在视觉上判定他面前的东西便是我们一致认为在他面前的东西，我们是不大会责怪他犯了视觉错误的。那么，我们会不会说他误解了指令，误解了被认为旨在告诉他看的是什么东西的指令呢？如果是这样的话，为不曲解结果，我们如何修正指令、以防止这种"误解"呢？要求他不利用先前的经验、避免一切概念把握，显然将把他置于无话可说的境地；因为，只要谈论，他就得使用语词。

① 哈克特出版社出版，1978 年。

把握“形式”与经验联想

古代的柏拉图主义者或新柏拉图主义者对这个问题的处理会简单得多。这样的哲学家会说，当我们注意一特殊感觉时，我们也就领会了一个“共相”或一个“形式”，也就是说，心灵具有把握性质自身而不仅仅注意显示这些性质的实例的本领。这样的哲学家会说，维特根斯坦和古德曼之所以觉得相似论有问题，是因为他们持唯名论观点，因为他们拒绝同“形式”、同对形式的直接把握，发生任何联系。

仅仅设想一种“把握形式”的神秘能力很难称得上一个答案；可是我们确实有一种与这种能力类似的东西。事物的性质确实是构成因果说明的一部分；当我们有一个感觉，而这个感觉又引出“这是对红的感觉”这样一种反应时，引起我的反应的部分原因是此感觉有一性质这个事实。确实，有些哲学家十足的唯名论立场使他们根本否认有这种作为“性质”的实体；但是科学本身对谈论性质是毫无顾忌的。我们难道不能说，当维特根斯坦的私语者(即企图发明一种私人语言的那个人)注意 X 并说出“E”时，“E”这个反应的原因是一种包含某种性质的因果关系，而且那种性质(不管它是什么)便是其他感觉要被正确归入 E 类便必须具有的与 X 的那种相关的“相似性”吗？

认为谈论性质在科学上完全允许的观点是正确的；但这无助于复兴柏拉图主义。我们同性质的相互作用只能通过同它们的实例的相互作用；而这些实例同时永远是许多性质的实例。不存在与一性质“自身”相互作用这类事情。谈论与一感觉有因果关系的性质并不能起到柏拉图主义哲学中那个作为感觉之(唯一)“形式”的概念所起的作用。

说得明白些，当我具有一种对蓝色的感觉时，我具有的是对蓝

色的感觉，同时我也有一种性质复杂的感觉，这种复杂的性质便是我之所以在那时把该感觉归于那个特定名称之下的原因。仅仅注意这个感觉还不构成对这些性质之一的“把握”。要把仅以这些方式之一而同我的感觉或同那个名称相联系的那个性质找出来，这个问题和我们极有缘分，即仍然是那个“恰当类型的因果之链”的问题。

为看出这点，我们首先要注意，当我们整个知觉经验引出“我正具有对蓝色的感觉”的反应时，我并非一贯正确。我本人就有过这种经历，在别人告诉我一件运动衫是绿色的之前，我曾三番两次称穿此运动衫的人为“穿蓝运动衫的那个人”。我并不是说那时运动衫看上去是蓝色的；在别人告诉我的那一刻我意识到对这运动衫作了错误的描绘。（我并不是经常有机会说“我正具有对蓝的感觉”，但如果我有机会这样说的话，那么在上述情况下，我可能在别人——他或许搞不懂为什么我在看着一个明明是绿的东西的时候会具有对蓝的感觉——责问我之前已经三番两次说了这样的话，而在别人问了之后，我会收回我刚才的现象报告。）这已经表明，引出“我正具有对蓝的感觉”这个报告的性质，或引出不管什么报告的性质，同成为对蓝的感觉的性质，或成为对不管什么有关性质的感觉的性质，并不是同一种性质。

对于上述情形，哲学家经常称之为一种“口误”。但这在我看来却是一个不幸的术语。“绿的”这个词可能曾在我唇边，而我可能沮丧地发现自己正在说“蓝的”。那倒会是一种口误。但在我刚才描绘的情形中，直到别人对我的报告提疑问我才注意到我在做错误的描绘（否则的话，我可能永远意识不到）。

有人提出的另一种说明认为，当我说“蓝”时，我的意思是指绿。但现在应当明白，当我们谈论事物时，我们并不是把“意思”放在心里，而绕个弯子去意指事物。说我“意思是指”绿色，无非是说我立刻接受了那个更正（并且在意识到自己一直在说的那话时感

到好笑)。这无非是重复而不是说明所发生的事情。

不管该作什么说明(或许是我大脑的词语加工元件出了故障),重要的在于,正如几页以前描绘的A-H这个性质即使在环境并未出现马匹的情况下也会引出"我面前有一匹马"的报告一样,我的整个思维结构也有一种复杂的性质,当我并不具有对蓝的感觉时,也会引出"我正有一种对蓝的感觉"(要不,至少,如果我被人质疑的话,会否认我有过那种感觉)。经验联想的机制总是不完善的。如果我们打算规定只要我具有一个引起"我正有一种对蓝的感觉"的报告(或不但引起这报告,而且使得我进一步考虑后也不觉得这报告"错")的感觉,我就正有一种对蓝的感觉,那么,根据日常的心理学理论,或许根据科学的心理学理论也一样,就可能有这样的场合,即根据上述标准说我正有一种对蓝的感觉是真的,尽管由于种种原因之一,该感觉的性质并不是蓝的。不仅如此,正如维特根斯坦指出的,根据这样一个标准,不论什么在我看来是对的东西,都将是对的——也就是说,作出一个确实正确的有关我的感觉的报告和作出一个在我看来是正确的报告之间的区别将被放弃。或许我们应当放弃或至少限制这种区别;或许,正如古德曼似乎暗示的那样,是否"实在地"具有他以为他正具有的那种感觉,这个问题是毫无意义的,除非在一些特殊场合,例如某人如果被质问的话就收回那个报告;但是,一个形而上学实在论者是不可能采取放弃这个区别的步骤的,因为在实在状况与某人所判定的状况之间的尖锐区别,正是形而上学实在论的要旨。

对于过去感觉的性质会不会总是搞错?

要澄清其中的奥秘还可以用另一种方式,即考虑以下问题"对于过去感觉的性质会不会总是搞错?"根据相似说,答案显然是"会的"。因为,根据该理论,我以前的感觉与我现在用"对红的感

觉”、“疼痛”等等标签来描绘的感觉要么相似，要么不相似，而它们是否相似这个问题，同我当时是否把它们归于同样这些标签之下的问题，截然不同。或许这世界就是这样，我们在公元以后的偶数时刻所称的“对红的感觉”，其实在质上相似于我们在奇数时刻所称的“对绿的感觉”，但我们的记忆老是欺骗我们，使我们对此从未察觉。那样一来，我在一分钟以前归于“对红的感觉”的标签之下的感觉与我此刻归于这同一标签之下的感觉，可能并不相似。

但是，在这个所谓的可能性中，有某种奇怪的成分。单说一事。我们说“我从未察觉”时的意义是很强烈的：如果我把不同时候我的“对红的感觉”当作是关于各种相关的物理事件的可靠信号（比如火焰，这个不可靠近的信号，等等），那么我会在自己的所有行动中都获得成功。“错的”相似性类（即我用来统称我称为对红的感觉的感觉——而不管它们“实际上”并非都有相同“性质”——的那个类）就可能是在我的解题活动中最好还是加以运用的类。但这样一来，这难道还真是错误的相似性质吗？

如果我们并不假定相似性的概念是不言而喻的，那么上述事例可以被重新描绘成这样一个事例，即正告诉我们情形如何的外部观察者所谓的“相似性”的关系与我们所谓的“相似性”的关系根本不同。如果我们采取这样一种观点，那么说我们“实际上”搞错了以往感觉的假说就站不住脚了：从内在论的观点来看，离开我们的合理可接受性标准而在不同时候“相似”的感觉这样一个概念，是不可理解的。

再论真理的符合论

到现在为止，读者可能已坚信指称的相似论是完全没有生命力的了。但是，为什么我们要断定真理的符合论必须放弃呢？即使我们的概念与它们所指称的东西之间有“相似性”的想法无法成

立，难道不可能有某种抽象的同构性，或如果没有完整意义上的同构性，难道不可能有概念向（独立于心灵的）世界上的事物的某种抽象的投影？难道不可以把真理定义为这样一种同构性或投影吗？

这种建议的麻烦并不在于词或概念与其他实体之间不存在符合，而在于这种符合太多了。要挑出词或心理符号与独立于心灵的事物之间的仅仅一种符合，我们事先总得指称那些不依赖于心灵的事物。要选出一种两物之间的符合，不能仅仅把其中的一个抓得紧紧（或只对其中的一个采取任何其他行动）；要选出我们的概念与假定有的本体对象之间的一种符合，不接近这本体对象是不行的。

要看出这点，可采用以下方式。有时，在不相容的理论之间实际上也可以互相转译。例如，如果牛顿理论是真的，那么每一单个物理事件可以按如下两种方式进行描述：或者是用穿过虚无空间发生超距作用的粒子进行描述（牛顿对作为作用力的引力就是这样描述的），或者用作用于场的粒子来进行描述，而这种场又作用于其他场（或同一场的另外部分），因此这种粒子对其他粒子发生的最终是一种接触作用。例如，描述电磁场行为的麦克斯韦方程在数学上等价于这样一种理论，在这一理论中，粒子之间只存在超距作用，这些粒子同时按照平方反比定律吸引、排斥，它们的运动并非不需时间，而是按光速（“延迟势”）进行。从形而上学的观点来看，麦克斯韦理论与延迟势理论是不相容的，因为传递分离粒子互相作用的因果载体（“场”）要么存在，要么不存在（一个实在论者会这样说）。但这两个理论在数学上是可互相转译的。因此，假如存在着一种使其中之一为真的同本体之物的“符合”，就可以界定使另一理论为真的另一种符合。如果使得一理论为真的不外乎抽象的符合（不管它是哪一种符合），那么几种不相容理论就都可能是真的。

对于内在论者来说，这是无可非议的。为什么有时不可以存在

几种同样好地与我们的经验信念相契合、同样融贯但又不相容的概念框架呢?如果真并不是(唯一的)符合,那么就展现了某种多元论的可能性。但是形而上学实在论者的动机是要拯救上帝的眼光,也就是说,拯救"唯一的真理论"。

不仅在对象和(我们认为)不相容的理论之间存在着符合(亦即,用逻辑学家的话来说,同一些对象可以成为若干理论的一个"模型"),而且,即使我们确定了理论并且确定了对象,这同一些对象可用来构成一既定理论的模型的不同方式也有无穷无尽(如果对象的数目是无限的话)。这不过是用数学语言陈述了以下直观事实,即要选取两个数域之间的对应,对这两个数域必须作互相独立的逼近。

我们看到的是一种持续了两千多年的理论的遗产。这理论如此长命,形式如此繁多(尽管从一开始就出现了内部矛盾和含糊不清),这个事实充分表明,对上帝的眼光的渴求是多么自然,多么强烈。康德第一次告诉我们这种渴求是实现不了的,但他又认为这种渴求根植于我们理性的本性自身之中(他主张在一个由社会体制和个人关系构成的完善体系中调和道德秩序和经验秩序,以此来设法实现"世界的至善",并建议在这一努力中对上述"总体性的"冲动予以升华)。这种自然的但无法实现的冲动的仍然存在,或许便是泛滥于我们文化中的假一元论和假多元论的一个深层原因;尽管它可能如此,我们现在可以抛弃上帝的眼光了。

第四章　心和身

身心平行论，身心相互作用论，身心同一论

在17世纪，大哲学家笛卡儿、斯宾诺莎和莱布尼茨都意识到，有一个关于心灵和肉体之间关系的严重问题。当然，在某种程度上，这个关系对柏拉图以及他之后所有哲学家来说，就已经是一个问题了；但随着近代物理学的诞生，它更成了一个问题。在17世纪，人们开始意识到，物理世界令人惊奇地完全受因果性支配。它是怎么完全受因果性支配的，这一点牛顿物理学表述得最好：任何物体的运动都是某种力的作用结果。力可以完全由数来描述：三个数足以确定任何力的方向，一个数足以描述任何力的大小。一个力所造成的加速度的方向与该力的方向完全相同，而加速度的大小，则可以根据牛顿第一定律 F＝ma 从物体的质量和力的大小中推出。如果作用于一物体的力不只一种，则可以用力的平行四边形定律算出合力来。

重要的是承认，这样一种物理学，这样一种强调数和用于计算的精确代数的物理学，与中世纪的基本上是定性的思维相比，有多么大的差别。在中世纪的思想中，几乎任何东西都可以对任何别的东西施以“影响”（我们的 influenza 流行性感冒一词便是这种中世纪思维方式的遗迹。据认为，邪恶的精灵对空气施以影响——在意大利语中，这叫做 questa influenza——而空气则进一步影响病

人)。根据这样一种思维方式,说心灵能够“影响”肉体,就不那么耸人听闻了。

在我上文提到的那些哲学家的时代,数学的思维方式正渐渐出现,并把那种旧式的思维方式撇在一旁。虽然这种新思维方式直到牛顿才得到充分发展,但在一些具体问题上,笛卡儿已经想到了力的平行四边形,在一些更简单的问题上,利奥纳多·达·芬奇也已经想到了这一点。在这些思想家心目中物理学可以具有的研究方式,有点像今天它的研究方式。他们看出物理学所处理的是力和运动,并且反对那种定性的说明风格。相反,他们设想力学世界也有其本身的逻辑,我们或许可以说,它有一种“程序”,除非有什么东西干扰,否则它就一直按照这个程序运转下去。

在这些思想家看来,心理事件可能做以下两件事之一,(1)它们可能与物理事件(例如,大脑中的事件)互相平行,其模型是两只互相同步的时钟:身体是一只上了发条的时钟,不管是否快活,一直走到死亡,就像整个物理世界不管是否快活从创世走到末日审判(或用现代的说法,引力坍缩)一样。心理事件也走着自己或快活或倒运的道路,而且不知怎么的,或许是神意使然,它已被作了精心安排,以致大脑事件B总是恰好出现在感觉S出现的时候。(2)它们可能同物理世界发生相互作用。实际上,心理事件可能引起大脑事件,而大脑事件也可能引起心理事件。

笛卡儿那种声名狼藉的相互作用论观点,他的那个心灵能在物质非常非常稀薄时对之施加影响(实际上,它以某种方式推动松果腺中的物质)的主张,与其说是胡思乱想(看上去它似乎如此),倒不如说是一套中世纪信条的遗迹。[1] 根据这种早期的思维方式,心灵被认为施作用于“灵魂”,灵魂又进一步作用于“物

---

① C·S·刘易斯的《遗弃的形象》(尤其是第7章第6节)中对这种中世纪观点作了描绘(坎布里奇,1964年)。

质”，而对灵魂人们并不认为它是完全非物质的。中世纪哲学家倾向于在各种存在物系列中的任何两个相邻项之间引入一个中介物，而“灵魂”正是这种倾向所自然而然地促使他设定的那种中介材料。它好像是一种只有一小股冲力的气体。一旦驱逐了这种“灵魂”，并实际上认为心灵是完全非物质的——只有那时，心灵推动松果腺中即使非常稀薄的物质，才会显得非常奇怪。这是无法想象的。

相互作用论观点的最朴素形式把心灵看作是某种精灵，它能够依附在不同肉体(但据大量关于灵魂全新附体和“回忆前世”的流行书籍说，在它的思维、感觉、记忆和表现人格方面并无变化)，甚至能够在无肉体的情况下存在(依然进行思维、感觉、记忆和表现人格)。对这种几乎是迷信的相互作用论，人们很容易驳斥说：存在着大量证据(其中有些在17世纪已为人所知)表明，思维、感觉和记忆的功能根本离不开大脑。说实在的，根据这种相互作用论，我们为什么要长着复杂的大脑这一点，是完全莫明其妙的。如果只要一个“舵轮”就完全足够了，那么它会比人脑小得多。

为了回避这种科学的驳斥，像笛卡儿这样的老练的相互作用论者主张，心灵和大脑是一个基本统一体：在某种程度上，进行思维、感觉、记忆和表现人格的，正是心脑统一体。这意思是说，我们通常所说的心灵根本不是心灵，而是心脑统一体。然而，这个学说包含什么意思，主张某物可以包含两个像心灵和物质这样的据说非常不同的实体而仍然不失为一个基本统一体，这是什么意思，我们一点也不清楚。

另一种观点，即平行论的观点，也相当奇怪。什么东西使得心理事件与大脑事件相互伴随呢？17世纪有一位大胆的哲学家提出，心理事件同大脑事件和其他肉体事件其实可能是互相同一的，这位哲学家便是斯宾诺莎。以当代形式出现的这个主张认为，我在某个特定场合处于疼痛状态，可能相当于这样一个事件，即我的大

脑在该场合处于状态 B 这样一个事件。(为了表达这个观点,我还将说,根据这样一个观点,具有那种特定痛觉的性质,和处于状态 B 之中,两者是相互同一的。我倾向于这样谈论,因为我认为我们目前对性质的逻辑理论要胜过对事件的逻辑理论,但我认为这个观点用两种方式都可以表达。用有关性质的术语来说,这个观点的意思是说人的那个性质,即那人体验到感觉 Q,与处于大脑状态 B 这个性质,可能是同样的性质。)以这种方式提出这个主张的人当中有 18 世纪的狄德罗;在 19 世纪 40 年代和 50 年代,这个主张成了"主流"。人们第一次开始对唯物主义和身心同一论加以认真对待,并开始提出,像斯宾诺莎提出的那一类观点(或斯宾诺莎的观点减去它的精巧的神学及形而上学装饰)是正确的:我们实际上正在研究的只是一个世界;我们若要知道具有痛觉、听觉、视觉等等的状态其实是大脑状态,就必须从事大量科学研究,这个事实并不意味着它们不可能是大脑状态。

以当代形式出现的这种身心同一论最早是由一批作者提出的,其中最著名的一位便是澳大利亚哲学家 J·J·斯马特。起初这主张认为,一个感觉,比方说一种特定的、对蓝的感觉,与某种神经生理状态是互相同一的。这种主张的一种变体——我想它是由我首先提出的——叫做功能主义。[①] 在功能主义的观点看来,这儿确实存在着一种同一性,但斯马特在确立这同一关系中的另一项时,找错了一种大脑性质。在功能主义者看来,大脑具有某些在某种意义上讲并非物理的性质。

那么,我说大脑具有非物理性质,这是什么意思呢?我指的是可使用那些不提及脑物理学和脑化学的术语来定义的那些性质。

---

① N·布洛克的《心理学的哲学阅读文选》(哈佛,1980 年)收入了一组讨论功能主义的出色文章。我本人的论文刊在我的《心灵、语言和实在——哲学论文集》的第 2 卷(剑桥,1975 年),作为该书的第 14 至 22 章。

假如一个物理系统居然会有非物理性质这一点看起来很奇怪，那就来看看一台计算机吧。计算机拥有许多物理性质。比方说，它有某个重量；它有若干集成电路块，或诸如此类的东西。它有一些经济性质，例如有一个价格；它还有一些功能性质，例如有某个程序。这最后一种性质，便是非物理的，之所以这么说，是因为它可以被另一种不管其形而上学或本体论的成分组织实际上可能是什么的系统所实现。一个游魂可以表现某种程序，一个大脑可以表现某种程序，一台机器也可以表现某种程序，这三者即使在原料、材料方面截然不同，它们的功能状态也可以完全相同。

心理性质也显示出同样的特点；同一种心理性质（例如发怒）可以是在物理构造和化学构造上区别颇大的数千种物种（其中有些可能是外星生物，或许机器人哪天也会发怒）的成员的一种性质。功能主义者主张，20世纪最站得住脚的、不把心灵与物质当作两种独立实体的"一元的"理论，是把心理性质等同于功能性质的理论。

我至今仍然倾向于认为这个理论是正确的；或至少它是对心灵/肉体关系的一种正确的自然主义描绘。另外也有一些对这个关系的"唯灵论"的描绘，它们也是正确的，但不可归结为我们称为"自然"的那幅世界图景（说实在的，"合理性"、"真理"和"指称"这些概念便属于这种"唯灵论的"描述）。对此我将在后面（第六章）作些讨论。但这个事实并未使我泄气；因为，正如纳尔森·古德曼曾强调的那样，非实在论的一个吸引人之处就在于它承认有可能对世界作不同的正确描述。但我却倾向于这样的观点，即自然主义的描述是一个正确的描述，根据这种描述，思想形式、映象、感觉等等，是具有功能特点的物理事件；而我想在此处讨论的是我前几年发现的功能主义理论的一个困难，即，此理论在感觉的质的特征方面存在着困难。当人们想到一些相对来说比较抽象纯粹的心理状态（比方说我们所谓的"括号内"信念，亦即，一种仅从其"概念"内

容上加以考虑的思想），或像嫉妒或愤怒这样的范围不确定的情绪状态时，这些感受与整个系统的功能状态的互相同一看起来是完全可接受的；但是当人们想到具有一种当下显示的性质，例如，体验到蓝色的一种特殊阴影，上述同一就不可接受了。

我在讲课时用了多年的一个例子是对那个“色谱颠倒”的著名例子加以改变而成的。这个色谱颠倒的例子（见于洛克著作[①]）说的是一个小伙子，他边走边看，从而把蓝的看成红的，把红的看成蓝的（换言之，他的主观的颜色同彩色负片上的而非彩色正片上的颜色相像）。人们听到这个事例时的初步反应可能是说：“可怜的家伙，人们会同情并原谅他的。”但谁能知道呢？当他看到任何蓝色之物时，他觉得是红的，然而他从小已经学会把那颜色叫做“蓝色”，因此如果有人问他，那对象是什么颜色的，他会说：“蓝色的”。这样，谁也无法知道。

我修改之后的例子是这样的：请想象你的色谱在你一生的某一特定时刻发生了颠倒，而你**记得它从前是什么样的**。不存在关于“证实”的认识论问题。一天早晨你一觉醒来发现天空是红的，而你的红毛衣则变成蓝的了，每个人的脸色都变得非常可怕，就像在一张彩色负片上那样。哦，老天！或许，你能学着改变你的谈话方式，把你觉得是红的东西叫做“蓝的”，或许你能充分调节到这样的地步，如果有人问你，某人的毛衣是什么颜色，你会给出“正常的”回答。但到了夜里，我们想象一下你会悲叹道：“哦，但愿颜色看上去还像我小时候看到的那样；颜色看上去竟与从前不同了。”

在这个例子中，人们甚至好像已经知道一定发生了什么事情。大脑中一定有一些“导线”发生了“混线”。来自蓝光的输入原来进入大脑中的这一个机构，现在却进入了另一个机构；而来自红光的输入则进入了前一个机构。换而言之，某种东西改换了**显像**，或物

① 《人类理解论》，第2卷，第32章（第14节）。

理状态。从前起着用信号报告环境中出现了"客观的"蓝色的功能作用的那个物理状态，现在则用信号报告环境中出现了"客观的"红色。

现在假定我们对主观的颜色采纳以下"功能主义的"理论："当一感觉(或大脑中与之相应的物理事件)具有用信号报告环境中出现了客观的蓝色的作用时，这感觉是一种对蓝色的感觉(亦即，它具有我现在按那个方式描述的那种质的特征)。"这个理论把握住的只是"对蓝色的感觉"这个词组的一个意思，而不是想要把握的那种"质的"意思。假如这种功能作用等同于质的特征，那么人们就不可能说那感觉的质业已改变。(如果这还不够明白的话，那么请想象在你发生色谱颠倒以后，在你学会在语言上予以补偿之后，你不幸突然得了健忘症，它抹去了关于从前颜色是什么样子的全部记忆。在这种情况下，你现在叫作"对蓝色的感觉"的那种感觉看来似乎可能和你以前叫作"对蓝色的感觉"的那种感觉具有几乎完全一样的功能性作用，然而，它所具有的特征却完全不同。)然而质确已改变。在这个例子中，质看来并不是一种功能性的状态。

在我看来，如果上述情况真可能发生，那么对功能主义者来说最可行的步骤是说："正是这样，然而，质的特征，只不过是物理的显像。"并且说，就这种特殊的心理性质而言，就质而言，那旧式的身心同一论便是正确的。如果读者非常倾向于唯物主义，他或她就可能认为"具有感觉"这个性质是一种大脑性质。不倾向于唯物主义的读者可能会认为"具有感觉"这个性质与一种大脑状态是相关的。或许大多数人都持有以下两种观点中的一种：认为感觉状态与大脑状态相关的观点，或认为感觉状态与大脑状态同一的观点。这个问题常常一再地以同样的方式引起争论。讨论老是采取这样的方式："假定 B 与 Q 相关，B 到底和 Q 同一吗？"我们知道这种感觉状态与这种大脑状态相互平行，"感觉状态与大脑状态互相同一"

这个说法对不对?讨论越是以这种方式进行,“相关”这个概念就越显得不成问题。对相关人们之所以不(大)讨论,是因为人人都知道至少存在着一种相关。同一之所以引起讨论,是因为成问题的正是它。但我将要设法表明的是,即使“相关”也是成问题的,不仅是因为存在着“不相关”的证据,而且因为从认识论上讲,即使存在着相关,人们对它是什么也无从知晓。问题的关键不在于主张唯物主义,而在于这个事实:我们认为至少存在着一种相关。

身心同一论和先验性

人们之所以有可能重新对身心同一论和其他“一元的”理论发生兴趣,是因为认识论气氛发生了变化,如果这不适用于重新发生兴趣的起初阶段(即在斯马特和早先的一些同一论者那里),至少适用于开始于 1960 年的这种情况。在 20 世纪 60 年代以前,人们对同一论并不认真对待,其理由是哲学家们“明白”它是假的,而他们认为自己明白它是假的,并不是基于经验的论据(因为,什么样的经验证据能表明一感觉状态不是一大脑状态呢?),而是基于先验的证据。人们考虑这个理论,一下子先验地看出一感觉状态不能是一大脑状态,也就是说,人们或许是看出,说感觉状态是大脑状态毫无意义,就像说数字 3 是蓝色的毫无意义一样。在 1950 年或 1960 年以前,人们以为他们一下便知道,或许多人以为他们知道,感觉状态不可能是物理的,另一些人认为这些人错了,但要争辩是无济于事的,多数派会说:“瞧,我们不能向您证明一感觉状态不可能是一个神经生理状态,我们不能向您证明每个数字皆有一个后继数,我们不能向您证明数字 3 不是蓝色的,但对这些事我们一看即知;它们是理性的真理。说一感觉状态是一神经生理状态,是无意义的、不可能的,对这一点我们再明白不过了。”这样,知道感觉状态不可能是大脑状态的是多数派,而知道多数派错了的则是少

数派。两派都先验地知道对方是错的。设想辩论的可能性，或走出争论的这种僵局的可能性，不大可能。

1951年，W·V·蒯因发表了一篇题为"经验论的两个教条"[①]的论文。从那时起，对"先验真理"这个概念的哲学信念便不断消融。蒯因指出，我们认为我们先验地知道的许多事情都不得不作了修改。这样，请看下例：假定有人向欧几里得提出，可能出现这样的情况，即我们可以有两条同第三条直线垂直而它们之间则相交的直线。欧几里得会说，这种情况不可能出现，这一点是必然为真的。但是根据我们今天接受的物理学理论，这种情况确实发生了。经过太阳附近的光线所表现的那种行为，并不是因为光沿曲线运动，而是因为光继续沿直线运动，而直线在我们的非欧世界中则具有那样的特征。

一旦我们承认以上说法，有的哲学家就一定要提出这样一个问题："还有什么先验性可言呢？"蒯因就提出了这个问题。（蒯因还令人信服地表明，对于先验性的标准的经验论解释——例如，"根据约定为真"的概念——是逻辑上非连贯的，但我不打算评论他的论据。）

但我认为，蒯因在某些方面是走得太远了。蒯因的"任何陈述都无法避免修正"的论断暗示，对每一个陈述来说，都存在着这样的场合，在这种场合下对它加以拒绝是合理的。但这显而易见是假的：到底在何种场合下，拒绝"并非每个陈述是真"，亦即接受"所有陈述是真"，会是合理的呢？[②]

但是如果说蒯因在反对先验性方面确是言过其实的话，他的正确之处则在于：我们的合理性概念和合理可修正性概念，并不是

① "经验论的两个教条"首次发表于《哲学评论》，1951年，它重印于蒯因的《从逻辑的观点看》（纽约，1961年）。

② 我在"分析性和先验性：超越维特根斯坦和蒯因"（载于《中西部哲学研究》，第4卷，1979年，明尼苏达）中讨论了蒯因对先验性这一概念的抨击。

由某一套永恒不变的规则确定的，并不是铭刻在我们的超验本性之中的（就像康德认为的那样），其有力根据恰恰在于，关于一种超验的本性，一种我们在本体上具有的本性，一种同我们从历史的角度或生物学的角度可以进行的自我设想无关的那种本性的全部想法，是毫无意义的。我们的合理性概念和合理可接受性概念产生于我们局限性极大的经验和可错性极大的生物组织，既然如此，就可想而知即使是那些我们当作“先验的”或“概念的”或诸如此类的原则，也将由于出乎预料的经验或突如其来的理论发明而不时被证明为是需要修正的，这种修正不可能是无限制的；要不然，我们就不再会有关于任何我们能称为“合理性”的东西的概念；但是这限制一般来说我们是无法说出的。除了一些琐碎的例子（例如，“并非每个陈述为真”）之外，我们无法担保，放弃一个现在被当作（而且，在一特定情境之中，被合情合理地当作）“必然”真理的陈述，在任何情境下都决不会是合理的。笼统地说，我们必须承认，对简单性、总体效用和可接受性的考虑，会使我们放弃一些从前被当作先验的东西，而这是合理的。哲学已变得反先验论了。但是一旦我们承认了我们当作先验真理的东西具有依赖于情境的、相对的性质，我们也就放弃了反对身心同一性的那个唯一有力的“论据”。同一论者势必要指出这点，他们确实也曾提出了这点。因此我们便说形势有了变化。

我一直在使用“性质”这个概念，但我觉得，至少有两个“性质”概念，而在我们心目中它们却被搅在一起了。一个是非常古老的用“谓词”这个术语来表达的概念（例如，在“存在是一个谓词吗？”这个著名问题之中），而另一个是我们今天在谈到“物理性质”、“量纲”等等时使用的概念。当一个哲学家心目中存在的是那个古老的概念时，他常常认为谈论性质同谈论概念是可以互换的。对这种哲学家来说，除非有一个说性质是相同的这样的概念真理，否则性质

就不可能是相同的；尤其是“具有一个带某种质的特征的感觉”这个性质，“同处于某种大脑状态”这个性质，不可能是相同的，因为两者对应的谓词不是(就广义的“分析地等价的”而言)同义的，而谓词的个体化原则恰恰是：只有当“是P”与“是Q”同义的时候，是P与是Q才是同一个谓词。

但是，请考虑，如果一位科学家断言温度就是平均分子动能，会出现什么情形。从表面上看，这是一个关于性质的同一性的陈述。他所断言的实际上(在“实际上”一词的某种意义上)是具有某一特定温度这个性质与具有某种分子能这个性质是同一个性质；换言之(更一般地说)，温度这个物理量与平均分子动能根本就是同一个物理量。如果这是正确的话，那么，既然“X有多少多少温度”与“X有如此这般平均分子动能”并不是同义的(即使“如此这般”是与温度的“多少多少”值相对应的分子能量值)，那位物理学家说“物理量”时所指的，同哲学家们所说的“谓词”或“概念”，也一定完全是两码事。

说得更具体些，区别在于，表达式“X是P”和“X是Q”的同义性对于谓词P和Q两者的“同一”来说是必要条件，对于性质P和Q两者相同来说就不是必要的了。与谓词相反，性质可以是“综合地同一的”。

如果“综合的性质同一性”确有其事，那么，处于某种大脑状态这个性质和具有一种带某种质的特征的感觉这个性质，为什么就不可以是同一个性质(同斯宾诺莎的思路完全合拍)——即使这两者的同一并不是一个概念真理，即使事实上在许多人看来这个说法是先验地假的——呢？人们提出的论证正是如此。简言之，一种反先验论浪潮，一种综合的性质同一性的新机制，这两者使得同一论者，尤其功能主义者好像不知不觉地引起重视了。

下面我想考察一下，如果在以上两者之外再加上第三者，会出现什么情形。如果一位哲学家(1)是反先验论的自然主义者，(2)承认有综合的性质同一性这类事，并且(3)持有一种强硬的实在论真理观，会出现什么样的情形呢？我想要指出的是，这样一位哲学家会遇到一些严重的认识论困难。

裂　　脑

我们来看看神经学家们最近20年中从事的一项特殊实验。这就是著名的"裂脑"，即大脑分解实验。我想讨论一下这种实验与同一论以及在整个讨论中被视为理所当然的观点即存在着一种相关性的观点，是不是有关系。

根据把大脑当作一个与计算机相似的认知系统的大脑模型，大脑具有一种语言，一种内部语言(可以是内在的，也可以是由一种内部"语言"或表征系统加上一种公共语言所构成的混物体)。有些哲学家甚至为这种假想的大脑语言起了一个名字，叫做"心语"。我们来看看，当根据这样的模型某人有一个视觉时(我将凭空虚构我的神经学，因为我了解得不够，但我想没有人是真正充分了解的)，会出现什么情形。以下便是一可能发生的故事：

当某人有一个感觉时，就作出一个"判断"；大脑必须"铭记"诸如"红色呈现于12点钟"之类的事情。所以，这个质(称为"Q")首先便对应于一个心语记录。此外，言语加工中枢也得到一个输入，这个中枢同喉相联，大脑之所以能用公共语言报告"现在出现红色"，就是因此之故。在语言中枢得到一个输入之前，以心语作出的判断可能先得从一个区域传递到另一个区域。在视皮质中也有一些事件(神经学家休贝尔和威塞尔曾对此作过研究)，这些事件我想象是发生于通往"心语记录"的中途和言语加工过程的。这些"记录"、"输入"和其他事件可以发生于不同脑叶上：如果胼胝体分裂，

则此人的右脑叶(没有言语的脑叶)可能看到红色(或至少,对一份只可为该脑叶看出的书面问卷,它将发出肯定的信号),但如果有人问这位被试,那张卡片是什么颜色,他将回答:“我看不到这张卡片。”最后,在某点上还形成了一种记忆线索或几种记忆线索(可以把它分为短期记忆和长期记忆)。几乎可以肯定说这不是一种线性的因果之链,很可能存在着一些分化和聚合过程,存在着一种因果网络。

问题在于,心理学是以一种间断性很强的方式分析心理事件的。这里有一种蓝色的感觉。这下它开始了;这下它结束了。但因果网络并不是间断性的。并不存在着作为感觉的唯一相关物的一种独一无二的物理事件。

如果身心同一说是正确的,感觉状态Q就等同于这种或那种大脑状态。对于Q等同于哪一种大脑状态,形而上学实在论者不可能认为从任何意义上说是一件约定的或决定的事情,或是有一种约定的成分。他们的观点是,事实上我们生活于这样一个世界,在其中我们体验为感觉的一些质的特征,与我们以别的方式作为大脑事件的物质性质而遇到的一些性质,实际上完全是同一些性质。(或表达得更好一些,在这个世界中,“具有带某种质的特征的感觉”这个性质,与“处于某种大脑状态”这个性质,实际上正是一回事。)

我们且停一下看看这种观点到底说些什么。假定红色是那种我们正注意着的主观的质(比方说,是由于先注视一张绿色桌子然后移开桌子,得到一个余像而形成的)。假定当我经验到这种红色时,我所处的感觉状态等同于诸大脑状态的一个析取。它不可能等同于某一个极为专门化的大脑状态,因为我们知道,可以取走任何一个神经元,或诸如此类的东西,而仍然具有那种经验。但这性质可能是一个析取的性质,比方说(令人难以置信地),要么某某区域偶数编号的神经细胞激发,要么质数编号的神经细胞激发。其实,

它可能是一个比上述析取大得多的析取。有可能存在着一个由各种神经状态构成的巨大集合，从而这些神经状态的析取可能是“经验到红色”这个性质。

我们还得稍微再深入一些。假如某某区域中偶数编号的神经细胞激发，我经验到红色。假如测脑仪说：“不，是某某区域的质数编号的神经细胞在激发，”我仍然经验到红色。这就是说，我分不出我处于这些大脑状态的哪一个之中。如果我经验到红色，则我必须处于它们中的一个状态之中，但我无法对它们作出区别某某区域中偶数编号的神经细胞正在激发，这并不是一种可观察性质。即使知道同一论是真的，从我的感觉中我也说不出我具有这种性质。且把这种性质称为“P_1”，而把奇数编号的神经细胞在激发这个性质称为“P_2”。感觉到红色这种感觉状态等同于析取式(P_1 或 P_2)，此处这当然是第二种性质。P_1 不是感觉状态，P_2 也不是感觉状态，只有它们的析取才是一种感觉状态。换言之，根据这种本体论，两个本身不可观察的性质的析取可能是可观察的。我经验为一种简单所与的，正是若干不可观察性质的一种复杂的逻辑函项。这就是形而上学实在论的观点。

这个观点经我一说可能听起来有点愚蠢了。为此我的一个朋友曾指出：“假定我们具有的探测 μ-介子的唯一装置辨别不出μ-介子和反 μ-介子。μ-介子不是一种可观察性质，反 μ-介子也不是一种可观察性质，而它们的析取则是。对于那些不认为可观察性像它实际上所是的那样是一个实用性概念的人来说，这只会显得自相矛盾。”但是，我的目的并不是要嘲笑这个观点，这个观点实际上构成了一个非常重要、非常合理的神经生理学研究纲领；我的目的是要阐明它要人承认的是什么。我将论证，导致困难的并不是同一论本身，而是同一论与形而上学实在论的合取——亦即，同我所谓的对真理本质的“外在论”观点的合取。

不接受这样一个观点还是可能的。这样，卡尔纳普会说(至少

在某一时期），谈论物理对象就是经过好几重推导来谈论感觉，决定说一种特定的大脑状态等同于一种感觉状态 Q，实际上就是决定以某种方式修改谈论物理性质的语言，决定改变我们关于上述物理对象的概念。

既然谈论物理对象和物理性质就是经过好几重推导谈论感觉，我们可以对规则加以修改。但那个观点并不是形而上学实在论的观点，至少在物质对象和物理性质方面。抱这种想法的某人可以是关于感觉的形而上学实在论者，但不是关于物质对象的形而上学实在论者，而且既然他认为对物质对象的谈论有点口气不硬，他为了采纳同一说只需说“我把它当作一种约定而采纳，当作一种对意义的进一步规定而采纳”就可以了。既然意义不完全是预先确定的，既然存在着某种结构开放性，就不存在“你怎能知道感觉状态与这个性质而非别的性质相同一”这个问题。如果这个性质是什么有点含糊不清，我们还是有余地干脆把同一性规定为一种对意义的说明。但我正与之讨论的人确实认为外面存在着一个物质世界，而这并不是经过好几重推导来谈论感觉：他确实认为存在着物理性质；他还主张像“某某通道的神经元正在激发”这样的陈述断定了我们某些特定的物理性质，而这些性质与这种感觉状态要么同一，要么不同一。

这同样适用于丹尼尔·丹尼特这样的哲学家，他认为对感觉的谈论是相当含糊的，认为并不存在一个处于这种感觉状态、拥有一种带这种质的特征的感觉的定义明确的主观性质。我认为，他也可以把同一说当作一种意义规定来加以接受，这次它所规定的不是物质对象词项的意义，面是心理学词项的意义。然而，这仍然不是地道的形而上学实在论者的观点。

我所考察的是一位地道的实在论者，他认为：“对，我知道这种心理性质（感觉状态）是什么。我有过这种性质，我可以认出它来。我认为这是一种我所指的确定的心理性质。我知道 P_1 和 P_2 是什

么，从而也知道(P_1 或 P_2)是什么，知道感觉状态同它要么同一，要么不同一。"恰如一位天真的物理学家会说的那样："不存在约定的成分"(顺便说说，我想他可能错了)；"在决定温度是平均分子动能方面，不存在任何约定的成分，温度要么是平均分子动能，要么是别的什么性质。"这就是我所要考察的观点。

问题在于，如果采取了这种形而上学实在论的观点，那么，存在着的可能性比人们惯于考虑的要多得多。首先想到的可能性是，感觉状态同一于以下性质，即具有发生于视皮层的有关事件，并且具有以恰当方式用"心语"记载的记录，并且具有通往言语中心的输入，并且具有已形成的记忆线索——亦即，感觉状态被认为是同一于这种种性质的合取。然而，一旦我们考虑分解的可能性，那时我们对自己所要求的实际上是不是整个合取，就没有把握了。或许感觉只是视皮层中的事件？(亦即，具有感觉这个性质"实际上"就是具有发生于视皮层的事件这个性质。)

我们姑且假定确是这样。现在我们来假设，我们可以切断产生心语记录的那个过程，或至少切断通往言语中枢的输入。我们设想给被试看一张位于其视域左侧的红色卡片(这样此卡片就只"可被右脑叶所看见"，如神经学家们所说的)。然后，在右脑叶上会发生有关的视皮层事件，但是如果我们对这位被试说："你看见什么红的东西了吗？"这位被试将说："没有。"

好，根据我们用以确定某人有无感觉的一个标准，即真诚的言语报告的标准，我们就得说他不具有对红色的感觉，从而我们也驳倒了说Q(有关质的特征)同一于上述视皮层性质的那个理论。但有人会反对说："不，你根本没有驳倒那个理论。因为，这是一个什么样的观察者？那家伙的大脑节切成两半了。"就任何处于常态的观察者所能告诉的而言，Q同一于这种视皮层性质。而不处于常态的观察者则不作数。他们不可能作数。

困难在于，有一些身心同一论是观察上不可区别的，[①]这意思

是说，在所有处于常态的观察者的经验方面，这些理论导致相同的经验。

请考虑这个观点：除非具有通往言语中枢的输入，否则不会具有对红色的感觉。对此怎样加以证明或驳斥呢？可以认为，如果我们切开胼胝体，但有某种不通过言语处理元件的记忆的话，就有了一条出路；也就是说，我们先问那家伙他有没有对红色的感觉。他说："没有。"然后我们重新缝合胼胝体（如果你能做到的话，那将是一项绝技），并问："你刚才有对红色的感觉吗？"他会说："有的，但这事太奇怪了，要知道，我刚才有对红色的感觉，你问我有没有，我却听到自己真诚地告诉你我没有。"（我听说，病人们对这类情况通常是"自圆其说"或提供理由，而不是像我刚才想象的那样加以描述。）这样一个报告会表明，即使没有一种通往言语中枢的输入，刚才也有一种对红色的感觉吗？

不会。假如丹尼尔·丹尼特（他一度曾主张感觉就是通往言语中枢的输入，或持与此相近的观点[②]）想要协调这位被试的报告与他的理论，他必须说的只是："我并不否认后来发生了记得以前曾有过那感觉这样一个心理事件。我否认的是以前发生过那感觉。"根据这两种理论，那位被试后来都具有记得（不管是否正确）他以前有过红色感觉这种经验。

这里的分歧并不是虚构的。多数神经学家确实相信，"裂脑"

---

① "观察的不可区别性"这个概念是由克拉克·格列莫尔和大卫·马拉蒙特在《时-空理论的基础》（厄蒙，格列莫尔和施塔赫尔主编，明尼苏达科学哲学丛书第8卷，明尼苏达大学，1977年版）一书中讨论时-空理论的论文里提出的。时-空理论中的类似问题是，存在着一些这样的"可能"时-空（即相对论所允许的时空），它们在总的拓扑学性质上不同，但其中的观察者所具有的经验则可能完全相同。这样的例证常常被人用以下理由摒弃，即"简单性考虑"会告诉某人，他生活于其中的是哪一个时-空；但困难在于（正如马拉蒙特指出的），那个物理学理论（广义相对论）并没有说我们生活于符合它的规律的最简单的时-空之中。

② 丹尼特是在"建立一种关于意识的认知理论"一文中提出的，此文重印于他的《脑痒变》一书，布雷德福出版社，1978年。

病人的右脑叶是"有意识的"。实际上，这相当于说有时存在着一种对红色的感觉，或诸如此类的东西，即使并没有通往言语中枢的输入。(这观点常以这样的方式表达："存在着两个意识区。")但是，至少有一位著名的神经学家艾克尔斯主张，分离出来的右脑叶(或左脑叶，如果病人的言语中枢在右侧的话)是不具有意识的。根据艾克尔斯的观点，存在着一种一元的意识。他会说，那分离出来的右脑叶能"模仿"意识行为，并不表明这是又一个"意识区域"。

求助于像"选择较简单的那个理论"这样的方法论箴言也无济于事，因为看来并没有任何有关的"简单性"是为"一元"论所有而为"两区域"论所无的，或是为"两区域"论所有而为"一元"论所无的。在某一方面，或许"两区域"论是较简单的；它指出某些行为能力(它们为右脑叶所具有，即使它不具有言语)是足以表现出意识的，而这符合我们称动物(它们也不具有言语)为有意识的这样一个事实。但是在动物与一只未受损的大脑——其脑过程即使不包括言语，也仍然是统一的——之间，与裂脑的一片之间，存在着许多不同之处。如果这事例不是一个同我们有着如此密切关系的事例，如果我们不这样强烈地倾向于关于感觉的形而上学实在论，把这事例看作是一个要加以辩护、而不是克服的事例，难道会不符合我们最好的方法论直觉吗？

简言之，存在着若干观察上不可区别的身心同一论。如果同一论者对的话，看来就根本没有方法可以使他知道他在哪方面是对的；知道一特定感觉状态同哪个大脑状态是同一(或相关)的，对这样一个重要的观点作进一步阐明是完全值得的。

托马斯·内格尔[①] 曾提出过一个似乎有理的观点，认为人们不可能想象做一只蝙蝠会是什么样子。但这观点为什么会看起来

① "做一只蝙蝠会是什么样子？"，重印于N·布洛克《心理学的哲学阅读文选》(哈佛，1980年)。

有理呢？几年前我读唐纳德·格里芬写的一本关于蝙蝠的令人愉快的书。我这才意识到，蝙蝠与任何别的哺乳动物并没有根本的区别，我们事实上大都认为我们能够想象自己家的狗和猫有什么感觉。在蝙蝠方面，又有什么困难呢？

对了，蝙蝠能够听到的声音比我们能听到的声音高出几个八度。我不能想象做一只蝙蝠是什么样子，是指我不能想象回响定位感是什么。但这就一定那么困难吗？我曾经能听到比我在中年时能听到的声音高一个八度的声音。但是主观的音高并没有改变：在客观音高方面，我现在能听的最高音可能比我 10 岁时能听到的最高音低一个八度，但是我现在听到的那些最高音仍然具有接近音域极限（过此极限就不能听到）的声音总是使我感到尖利、刺耳的音质。或许，一种比我们听到的声音高五个八度的声音在蝙蝠听起来就是这样：像一种短促刺耳的高音。

现在让我们想象两位哲学家或心理学家之间的一场争论。其中一位说，蝙蝠的感觉性质与人们的感觉性质全无共同之处。这两者的感觉性质之间的差别大得令人难以想象。你永远也想象不出做一只蝙蝠（甚至一条狗或一只猫）会有什么感觉。我们可以想象，另一位哲学家则回答说："胡扯！或许确有一些蝙蝠的感觉我无法想象。有一些他人的感觉（例如，异性的某些感觉）我可能也无法想象，但是这并不意味着我把那些他人的心理空间与我自己的心理空间的区别看作大得令人难以想象。为什么我不可以认为例如蝙蝠的视域与我的视域非常相像呢？（注意，与民间传说的正相反，蝙蝠的视觉很好。）如果允许对蝙蝠眼睛的光学结构作某些调整，或使蝙蝠的听觉范围与我的重合，它的听觉就像我的听觉，它的痛觉就像我的痛觉。"这样，我们能解决这场争论吗？

因为神经元的数目不同，因为结构比例不同（蝙蝠大脑的听觉中枢达该脑的八分之七之多），所以，那些如果同一说正确的话便与蝙蝠的感觉性质同一的、处于最最专门化的神经学层次（那儿有

大量神经元激发着)的性质,不可能就是那些与人类任何感受性质相同一的性质。它们可能吗?假定一只蝙蝠具有某种视觉(由于看见红色对象而产生的),假定该蝙蝠具有析取性质(P_1或者P_2),其中P_1和P_2是该蝙蝠的极为专门化的状态(实际上它可能是一个包括成千上万情况的复杂得多的析取性质,但我们予以简化吧),我们来这样假定,当我具有某种(由于看见红色对象而产生的)视觉时,我的大脑具有析取性质(P_1'或P_2')。考虑以下两个理论:(1)那只蝙蝠的感觉(称之为"红$_B$")的质的特征同一于(或至少相关于)析取性质(P_1或P_2),而人类感觉(称之为"红$_H$")同一于(或至少相关于)另一个性质(P_1'或P_2')。(2)那只蝙蝠的感觉的质的特征同一于我的感觉的质的特征(即,红$_B$=红$_H$),而两者同一于(或至少相关于)更复杂的析取性质(P_1或P_2或P_1'或P_2')。

根据第一种理论,蝙蝠和我有不同的经验,而根据第二种理论,我们具有相同的经验;但是,在人类观察者(不管是正常的还是异常的)将经验到什么这一方面,这两种理论导致相同的预见。同以前一样,它们又是在观察上不可区别的。

一些方法论箴言(如"选择较简单的那个理论")能有所帮助吗?同以前一样,还看不出有这种迹象。内德·布洛克曾指出,在某一方面,第一种理论较简单(在多种情况下都同一于一种较简单的物理性质),但在另一方面,第二种理论更简单(第二种理论是"反沙文主义的",它承认,要具有我们的感觉性质,不一定非完全具有我们的物理构造不可)。因而,同以前一样,我们仍然缺少原则去确定一个唯一可取的权衡。实际上,有什么理由认为应当有或者必须有这样的原则呢?为什么我们不可以像维特根斯坦鼓励我们做的那样,放弃我们关于感觉和关于"同样"(在涉及感觉时)的形而上学实在论,并把这一事例也当作一个要加以辩护而不是加以克服的事例呢?

最后,我想摆出三种我断定是错误的理论,但如果形而上学实

在论正确的话它们便是难以排除或无法排除的。它们是：(1)红$_H$毕竟等同于一种功能性的(或准功能性的)状态，也就是说等同于这样的状态：处于在某人早年曾起过用信号正常地报告出现了客观的红色这个功能性作用的任何物质的(比方说物理的)状态。(2)岩石具有感觉性质(亦即，在质上可能与视觉相似的事件发生在岩石身上)。(3)民族具有意识。

我们先来考察(1)。读者还记得我用以表明红$_H$不可能是一种功能性状态的那个论证。那个论证说，如果我们把红$_H$等同于“处于用信号正常地报告客观的红色出现的任何物质状态(例如大脑状态)”这样一个功能性状态，那么，我就不可能有过一种色谱颠倒(至少在“健忘”的情况下)，因为在看见某种客观上是红色的东西时，不管先于色谱颠倒还是后于色谱颠倒(如果有时间进行语言上的调整的话，而且，如果有必要的话，设定患了健忘症)，我都处于这种功能性状态之中。但是，根据一位形而上学实在论者的观点，我经历过一场色谱颠倒当然是有可能的(即使我因为健忘症发作而不记得了)。如果我没有发作健忘症，并且想起来我的色谱已被颠倒，这事例就更有说服力了。即使在这样的情况下，如果已能自动进行语言调整，仍然可以说以前是“对绿色的感觉”的东西，现在却起“用信号报告环境中出现了客观的红色”这样的功能性作用。

这个论证只表明红$_H$不等同于刚才提到的功能性状态，而并不表明它不等同于一种更为复杂的功能状态，比如这样的状态：处于在某人早年实现上述功能状态的任何物质状态。有人可能会反对说，这是一种滑稽的性质，一种若干功能性质的复杂的逻辑函项。但是，为什么若干功能性质的一种复杂的逻辑函项不比复杂的物理性质的一种析取那样更有可能与红$_H$相同一呢？难道这世界倾向于物理性质的析取而不倾向于功能性质的合取吗？

我们来考虑(2)。令P_3为“成为岩石”这个性质，并考虑以下假

设：红$_H$等同于析取性质(P_1 或 P_2 或 P_1' 或 P_2' 或 P_3)。当然，岩石始终具有这种性质。所以，根据这个假说，带有红$_H$这个质的特征的事件始终发生于岩石之中。(它们并不是在功能意义上的经验到红色，但是，在岩石中始终发生着这样的事件，它带有"在我们身上起那种成为红色感觉的那种功能作用"的事件的质的特征。)或者考虑一个更为复杂的假说，根据这个假说，岩石在不同时候具有不同的感觉性质。或者干脆这个假说：这些假说(其细节不详)中的某一个是正确的。我们可能会说"这些假说是疯人的呓语"。是的，此话不假。但是其中每一个在所有人类观察者那里，都和那个"清醒的"理论导致同样的预见。哪一个也不能凭借观察的或实验的根据予以排除，因为从标准观点来看，其中每一个都是观察上不可区别的。

我们可能以为，为了排除这些理论，可以求助于以下方法论原则：不应该无理由地把一性质归于一对象。当然，这原则并没有说这些理论是假的(有时候，我们没有理由相信的东西，倒是真的)，但至少它指出我们把它们当作假的是有正当理由的。但是，果真没有理由坚持这些理论的那个最起码细节(认为某个这样的理论是正确的，并且岩石具有感觉性质)吗？对于以下论证，即如果我们具有感觉性质，而物理主义是真的(许多哲学家认为有好多理由接受物理主义)的话，至少有一个物理对象发生了带一种质的特征的事件，我们怎么看呢？为什么这样的事件不可以发生于所有物理对象之中呢？如果我们能够表明，感觉性质本身有某种东西要求它具有像它在人类的感觉性质那里所具有的那种特殊的功能"作用"，那么上述道路就堵塞了。但是这正是那些相信感觉性质是一种形而上地实在的对象的人们告诫我们不能接受的。

最后(但并非最不重要的)，我们来考虑一下(3)。请考虑以下假说：痛苦等同于有机体和民族都能够表现的一种合适的功能状态。换言之，假定美国宣布"美国因……感到痛苦"，它确实痛苦。当

然，我们可能永远也不知道。或许读者此时正觉得有趣，并稍微有点好笑，一个团体在宣布其痛苦时，其行为方式居然能够类似于确实感到痛苦的东西的行为方式；但是读者认为美国并非真地感到痛苦。而根据上述假说，那么这读者就错了，民族精神确实会感到痛苦。

这个假说与心灵哲学中的一场有趣的讨论有关。功能主义者们（包括本人）喜欢使用的一个论证是以下“反沙文主义”论证：原则上，一台机器人和一个人（不管怎么样，在功能组织方面）之间的差异可还原为物理组织和化学组织的微小细节。甚至可以有一台与我们一直小到神经元都互相符合的机器人。（除了大脑，它甚至还可以有一具“血肉”之躯。）区别在于，我们的神经元是碳、氢和蛋白质等等构成的，而机器人的神经元则是由电子元件构成的。但是，从神经元往上，所有的线路都完全符合。这样，除非你是一位认为碳和氢内在地更有意识的“氢碳沙文主义者”，你为什么不可以说这台机器人是一个其大脑里恰好金属较多而氢与碳较少的人呢？

这个论证引起了以下回答：“那好，咱们不谈按人类神经元的连接线路连接起来的电子配件、电子神经元，我们假定有一些微型人，一些小小的女童子军和男童子军。”我们甚至不必想象这些小人居然知道整个组织的目的，不必想象除了他们在其中互相传递信息的一间或许多间光线昏暗的房间之外还看见什么。（当然，与我们的时间相比，他们的时间可能流逝得相当快。）他们可能是异化了的工人。“好，”那人接着回答，“你不会因为你知道实际只是这些小人在推动机身就称那东西为‘有意识的’。而这表明，一个（同我们的功能组织相像的）合适的功能组织，不足以为运用像‘有意识的’这样的谓词提供依据。”

对上述回答的一个回答（也就是我实际上作出的回答）是否认“多头机器人”（对刚才提到的东西人们就是这样称呼的）真地具有

和我们一样的功能组织。但我本来还可以作一个更为激进的回答。我本来还可以说："为什么我们不可以把这多头机器人称为有意识的呢？如果第一个论证是正确的（我认为它确实正确），如果那台具有正电子大脑的机器人可以是有意识的，为什么多头机器人的神经元较有意识这个事实会意味着它们的整体却较无意识呢？不管怎么样，我们在某种意义上是一个由若干小动物构成的社会。我们的细胞在某种意义上说是一个个动物。或许它们稍微有点感觉，谁知道呢？在此之上，则是我们的感觉。"好，如果沿此方向推论，如果我们断定那多头机器人是有意识的（即使它的神经元是男童子军和女童子军），那么为什么美国就不具有意识呢？

当然，我并不声称美国具有和人类一样的功能组织。它显然不具有这样的组织。但是存在着许多相似之处。美国有防务组织。它有消化组织，它吃油和铜等等。它还大量排泄（污染）。在功能组织方面，它或许像一只来回飞舞的苍蝇（我们认为苍蝇确实是有痛感的）那样与哺乳动物相似，难道不是那样吗？

"质的特征"定义得是否明确？

行文至此，我们还没有怀疑过以下观点：说某人的两种感觉有没有同样的"质的特征"，这话的意义是一清二楚的。但是，即使在内省的层次上，它也不是一清二楚的。单说一事，即使某人的经验在他自己看来是什么，也依赖于先前的概念结构，这是众所周知的事实。比如，即使我们斜视一张圆桌，我们也报告说看到了一张圆桌面。

在这圆桌面的例子上，心理学家和哲学家们从 19 世纪起就开始争论：人们是具有"椭圆形感觉材料"，并且心想它们是圆的（除非某人是一个"训练有素的内省论者"），还是具有圆形的"格式塔"，并且心想由于光学理论的缘故，它们是椭圆形的。人们可以有

符合其中一个描述的经验；而许多经验将符合这两种描述。神经学也不大可能解决这场争论。视网膜上的椭圆视象无疑在大脑本身中产生了若干事件，如果我们把这些事件等同于“视觉”，那么我们很可能得出类似“椭圆形感觉材料加上无意识推论”这个经典说法；那经验（我看见一张圆桌面）的被判定的特征也符合于大脑中的“报告”和“输入”，而如果我们把这些报告和输入等同于视觉，我们很可能得出这样一个说法，根据这个说法，人们除非判断某物看上去是椭圆形的，否则就不具有椭圆形感觉材料。为什么我们不可以说这两种说法都是合情合理的呢？正像古德曼在谈到那位被要求描述视运动的被试时所说的：

> “我们最好对他要使用的那种术语、那种词汇加以具体说明，要他用知觉的或现象的术语，而不是物理学的术语，去描述他所看见的东西。不管这是否导致不同的反应，它都对于所发生的事件产生一种完全不同的影响。对用于塑造事实的工具必须加以具体说明，这一点就使得把物理的东西等同于实在的东西、而把知觉的东西等同于仅仅显现的东西，成为毫无意义的事情。物理事实不是知觉事实的一个完全人为的翻版，同样，知觉事实也不是物理事实的一个面目全非的翻版。[1]

如果我在同一天的两个不同时刻看见一块红色的桌布，我是具有相同的红色感觉呢，还是具有不同的感觉但没有注意到区别（除非我恰好是一个画家）呢？

一个特别使人困扰的事例是关于适应的事例。如果给一被试一副使物象上下颠倒的眼镜，过一段时间，被试眼里的东西又显得

---

① 《构造世界的种种方法》，第92—93页。

正常了。感觉材料是"突然转了回去"呢，还是被试熟悉了被改变了的感觉材料，并对"上"、"下"作了重新诠释呢？很可能这位被试自己也说不出，在哪一时刻事情变得正常了，或发生了这些事情中的哪些事情。（像我这样戴双光眼镜的读者可以自问：视域的下半部分即使不去注意它的区别它也会显得不同吗？）对有一些转换，被试永远也不能适应（事实上，人们适应的是一些比较简单的转换），我还曾经假定，人们不会适应颜色颠倒，尽管如此，适应现象无疑对"同样的质的特征"在多大程度上是一个明确定义的概念提出了疑问。

实在论和感觉性质

我们考虑了一组带怀疑性质的困难。这些困难要表明的并不是同一论是错的（或相关论是错的——请注意它们都可以被表述为"相关"论所面对的困难，就像表述为同一论的困难一样），而是：如果它是真的，就存在着大量具体说明细节的可选择方案，以至于决不可能知道其中何者为真。而不知道其中何者为真，意味着不知道用什么来回答大量传统的怀疑论疑问，例如，岩石和其他无生命对象有没有感觉性质，蝙蝠之类的物种有没有我们所具有的感觉性质，民族能不能感到痛苦，等等。

但是，为什么哲学家们居然认为，岩石可以有痛苦（即一种同人类的痛苦有相同的"质的"特征的事件可以发生于岩石"之中"）这甚至是一种逻辑上的可能性呢？或许罗素为我们提供了了解这种形而上学实在论的本质的某种线索。罗素是一位关于感觉性质的实在论者和一位关于共相的实在论者。此外，他把感觉性质当作一些范式性的共相。一个共相首先是事物间可能相似的一个方面；在罗素看来，"事物间可能相似的方面"在认识论上最为原始、最为基本的例证是人们自己的感觉的质的相似之处。对罗素来说，感觉

性质是典型的共相。

但是，在传统的实在论者看来，共相是完全明确定义的：词可能模糊，但共相本身却不可能模糊。（哥德尔在一次谈话中说，一个模糊的词之所以模糊，是因为它代表了一个模糊的概念集；但概念都是完全明确定义的。）

所以，如果感觉性质就是共相，而共相在本质上是明确定义的，那么，任何特定的事物或事件——包括裂脑的一半或其中的某种事件；包括一块岩石或其中的某种事件；包括一个民族或团体或其中的某种事件——有没有表现出任何特定的感觉性质，就必定是完全明确定义的。如果认为感觉性质独立于它所起的功能作用，如果认为以下事实是完全偶然的，即一个红色感觉的质的特征就是具有那种特定的功能作用的东西的质的特征，那么，裂脑或岩石具有那种感觉性质就确实显得是一种逻辑的可能性了。

对这些可能性中的每一个是否都有意义我想持否定态度（虽然其中有些可能有意义——确有一种把裂脑的右脑叶看成"意识区"的倾向，我曾指出，决定这样看待是正确的）；一个像我这样的哲学家必须费心去澄清他不是在信奉某种形式的行为主义。说"感觉性质"不是明确定义的实体，并不等于说它们不存在，不等于说只有行为，以及诸如此类的东西。许多概念虽然模糊，但仍然有一些明确的所指。例如"房子"这个概念，对"伊格鲁"（爱斯基摩人的圆顶小屋）来说（伊格鲁是房屋吗？），对"豪根"（美国西南部的印第安人的泥木建筑）来说，或许在其他情况下也一样，就是定义不明确。但是对伊格鲁是不是房屋这个问题没有搞明白，并不意味着房屋不存在。同样，右脑叶是不是"有意识"的问题没有搞明白，也并不意味着有意识的生物不存在。

在某种程度上，感觉的质的相似性是有明确意义的。如果先有一个对红色的感觉，后有一个对绿色的感觉，我知道我有过不同的感觉（我不比较两者的功能作用即可知道这一点），而如果我先有

一个对红色的感觉，后有一个“同样的”对红色的感觉，我知道（其模糊程度不超过我们上面讨论的）我有过相似的感觉。但是，对于持一种“内在论的”真理观的人来说，这并不意味着在两个感觉（更不用说两个任意事件）是否在质上相似的问题上，处处都存在一种既定事实。

令 E 为我在一特定时间具有一特定感觉这样一个事件，令 E′ 为岩石中的某种物理事件。说 E 的质的特征（比方说，红$_H$）与（P_1 或 P_2 或 P_3）这样的性质（其中 P_3 是“成为岩石”这个性质）可能同一或相关这个主张是违背任何健全的人类常识的。说 E 和 E′ 可能是在质上相似的事件这个主张是荒谬的。我们曾讨论过对这种荒谬性的一个说明：这假说之所以荒谬，是因为它违反了“不可毫无理由地将性质归于一对象”这个方法论箴言。即使这个说明能够成立，它可能导致的结论也远不是说岩石不可能具有*感觉性质*（或者说，认为它们具有感觉性质的观点是自相矛盾的）。如果认为岩石具有*感觉性质*的那个“假说”的全部错误就在这里，那么我们就可能不得不这样说：*说不定岩石可能是有感觉性质的，但这种情况的或然性先验地极低*。

事实上，说岩石具有感觉性质的假说之所以自相矛盾的情况和钵中之脑的假说如出一辙：像钵中之脑假说一样，这个“假说”也预设了一个神秘的指称论。在任何理智健全的人看来，E 和 E′都极不相似，以至于“质上相似性”（也就是说两个感觉可能在质上相似，亦即有同样感受）的问题甚至都不出现。但那位形而上学实在论者，虽然根本不否认这点，却甘冒被人视为“疯人呓语”之险，认为 E 和 E′可能（在逻辑上有可能）以这种方式相似。他持这种观点是因为他处于以下幻觉之中，以为具有了上述感觉，即带有其质的特征、其“感觉方式”、其功能作用以及相伴的思想和判断的上述感觉，他就以某种方式使得“这个感觉的感受方式”（或某种专用代入词，例如“这种感觉的质的特征”，或“红$_H$”，或“这种感觉性质”）这

样一个表达式指称一个确定的“共相”，指称若干形而上学的单个事件的一个绝对定义明确的性质。但事实并非如此。

如果真有一些在功能上与我们同态的机器人，我们与它们一起工作、与它们进行争论，与它们结成朋友，我们就会马上断定它们是有意识的。（对于它们和我们有没有同样的感觉性质，我们可能还感到迷惑不解；但我们不会经常想到这点，就像我们不会经常想到蝙蝠或狗和我们有没有同样的感觉性质一样。）但是，假定我们碰到了一些九头机器人。（设想它们实际上是在某个地方通过某种生物过程进化而来的，恰如处于共生关系的动物在地球上进化那样。）对这些机器人，我们会怎么想呢？

人们对这样希奇古怪的事例不可能真的有什么把握，但看来甚至在这儿（如果我们大多是与机器人整体打交道，而只是偶尔与它的有意识的“神经元”——我那个故事中的“男童子军和女童子军”——打交道的话），我们也开始把意识归之于它了；但或许我们始终都会有不同意见。如果我们终于断定那九头机器人是有意识的，那么，对于美国，我们会逐渐感到隐隐不安吗？我不得而知。

在上述所有事例中我所坚持的观点是：这儿并没有什么隐蔽的东西，并没有像一些实体真有意识或真没有意识这一类的本体事实，或者像一些性质实在地相同或实在地不同这一类的本体事实。存在着的只有如下明显的经验事实：岩石和民族与人类和动物大不一样；各种机器人介于各种对象之间；如此等等。岩石和民族并无意识；我们实际拥用的意识概念就是这么回事。

这条思路之所以显得令人不安，是因为它使得我们的合理可接受性标准、正当性标准，归根到底，真理的标准，依赖于显然产生于我们的生物遗产和文化遗产（例如，我们是否曾同“智能机器人”打过交道）的相似性标准。但是这类情况也适用于我们在日常生活中运用的大部分语言，大部分像“人”、“房子”、“雪”、“棕色”这样的词语。一个承认这个解答——对于感觉性质方面的一系列难题所

提出的这个解答——的实在论者可能会对它作这样的表述："感觉性质"并不实际存在，或感觉性质属于我们的"次阶概念系统"，但是把"房子"置于非存在一边的"存在"概念，有什么意义呢？我们的世界是一个人的世界，而什么是有意识的，什么是没有意识的，什么有感觉，什么没有感觉，什么同什么在质上彼此相似，什么则是不相似的，所有这些都最终依赖于我们人类对于相同和区别的判断。

第五章 对合理性的两种观念

在上一章我谈到了合理性，也谈到了“合理的可接受性”。但合理性可不是一件容易说清的事情。

类似的问题在别的一些领域中也有。几年前我研究过像黄金这样的自然物种词项的特征，结果我认为，该术语的外延不仅仅是由一组“语义规则”或其他惯例化规范所规定的。规范可能会规定某些对象是黄金的范例；但它们并不规定该术语的全部外延，而且，还有这样的可能，即一个范例被证明实际上并不是黄金（如果规范仅仅定义什么是黄金，就会出现这种情况）。

如果某物表明其具有与若干范例（或大多数范例）相同的根本性质（或更含糊地说，与它们有“足够的相似”）的话，即使我们现有的检验手段不足以表明它是该物种的一个成员，我们也准备把它算作是属于该物种的。什么是这种根本的性质，什么可当作是足够的相似性，既取决于自然物种，也取决于上下文（冰冻茶在某个上下文中可以是“水”，但在另一上下文中则不能是水）；但是对黄金来说，重要的是原始成分，因为从古希腊人起，人们就一直认为成分决定着实体的规律性行为特征。除非我们说古希腊人的“chrysos”（“金”、“金黄色”）的意思指的是同范例具有同样根本性质的任何东西，否则，他们对新的检测仿金物方法的探索（这种探索曾使阿基米德去进行那个密度试验）和他们在物理上的思辨，就都没有意义了。

我们很容易在合理性本身的问题上也采取同样的思路，并且

说，决定一个信念是否合理，并不是这个或那个文化的合理性规范，而是一个关于合理性的理想性理论，一个为某一信念在任何可能世界的相关场合成为合理信念提供必要条件和充分条件的理论。这样一个理论必须说明范例，就像一个关于黄金的理想性理论对黄金的范例提供说明一样；但它也可能不止于此，进而提供能使我们理解那些目前尚无法透彻理解的事例，就像目前关于黄金的理论能使我们理解最出色的古希腊人也不可能理解的事例一样。在把“合理的”、“有理的”、“被证明为正当的”等等当作自然物种术语方面，一个总的困难是真正发现关于一切合理的可接受信念的强有力概括，看来前景相当暗淡。所有黄金都遵守的严格的普遍规律是存在的，正是这种规律使我们有可能把黄金描述为这样的材料：当我们认识它的时候，它将被证明是遵守这些规律的；但是，我们有没有运气去发现一切被合理地证明为正当的信念所遵守的严格的普遍概括呢？

缺乏这方面的可能性，并不意味着在对黄金本性的科学研究和道德研究或哲学研究三者之间没有类似的情况。例如，在伦理学中，我们从关于一些个别行为是对的或是错的判断（也可以说是观察报告）入手，并逐渐建立起基于这些判断之上的公理（而不是毫无例外的概括），这些公理还常常伴随着像“将心比心，友善待人”（一种“低层次的概括”）这样的理由或例证。这些公理又进一步对我们关于个例的判断发生影响和进行修改，从而出现对先前的公理进行补充或修正的公理。在公理和关于个例的判断之间几千年的这种辩证之后，可能会有一位哲学家问世，并提出一个道德观（一个“理论”），它可能对公理和单个判断都加以改变，如此等等。

这完全同样的程序在所有哲学（其范围与合理性理论似乎相当）中都可以找到。在几年前的一篇文章[①]中我按格赖斯和贝克的方式描述了道德体系的必要成分。我把以下三个要求包括在

内：(1)至少人们的基本假定应当具有广泛的说服力；(2)人们的体系应当能够承受合理的批判；(3)所建议的道德应当是有活力的。

对合理性的本质形成更好理解的途径——我们知道的唯一途径——是按同样方式对合理性形成一些更好的哲学观念。(这是一个无休止的过程；但这也应当是一个无休止的过程。)我为一个道德体系列出的必要成分，如果居然能不加改变地列为任何与人类有关的重大领域中的方法论或合理程序体系的必要成分的话，那将是令人吃惊的。在分析哲学中，以这种方式更好地理解合理性之本质的主要努力，是由一些科学哲学家作出的，这些努力导致了两种重要倾向。

逻辑实证主义

在过去50年中，那股被称为"逻辑实证主义"的思潮最为明白地表达了这样一种倾向，即认为"合理的证明为正当"的方法是由一份清单或准则提出的，虽然人们都认为，科学哲学家在制定这份清单或准则方面并没有获得成功。实证主义者希望"科学逻辑学家(他们用这个术语称呼哲学家)某一天能成功地制定出来的这份清单或准则不仅要包罗无遗地描述"科学方法"；而且，既然在逻辑实证论者看来，科学方法就是合理性本身，而用这个方法进行检验的可能性就是有意义性("一个句子的意义就是它的证实方法")，所以，这份表格或准则将决定一个陈述是有认知意义的，还是没有认知意义的。可以用这份清单中的方法(数学方法、逻辑方法和经验科学方法)加以检验的陈述将被算作是有意义的；所有其他陈述，

---

① "文学，科学和反思"，《新文学史》，第7卷，1975—1976年，重印于拙著《意义和道德科学》，劳特利奇和基根·保罗出版公司，1978年。

实证主义则都认为是“伪陈述”，或伪装的废话。

一个明显的反驳是说逻辑实证论者的意义标准是自相反驳的：因为这标准本身既不是(a)“分析的”(这个术语实证论者用于解释逻辑和数学)，也不是(b)(经验)可检验的。最奇怪的是这个批判居然对逻辑实证论者触动甚微，对这个运动的发展也妨碍极小。我相信对这个特殊的哲学攻势的忽视是一大失策；这个攻势不但是正确的，它还包含了一个深刻的教训，一个不仅仅针对逻辑实证主义者的教训。

我将要发挥的观点建立在以下意见的基础上，逻辑实证论者所允许的那些“证实形式”，就是已被近代社会“惯例化”了的形式。在实证论意义上可以被“证实”的，依不同的情况，也可以被证实为正确的（按“正确”一词的非哲学的或前哲学的意义来说），或可能正确的，或高度成功的科学；而公众对正确性、或可能正确性、或“高度成功的科学理论”的地位的承认，则是对我们的文化所持有的知识形象和合理性规范的具体体现、热情颂扬和推波助澜。

从表面上来看，最初的实证主义的证实范式并不是上述成为公认惯例的范式。在卡尔纳普的《世界的逻辑构造》一书中，证实归根结底是私人的，是以感觉为基础的，而这种感觉的主观的质或“内容”则据说是“不可交流的”。但是，在纽拉特的再三鼓动下，卡尔纳普很快转向一种把证实看成是较为公认的、较为“主体间”的这样一种观点。

波普尔曾强调以下观点，即认为科学预见会碰到一些“基本语句”，也就是像“天平的右秤盘下垂”这样的语句，它们即使无法“证明”得使怀疑论者心悦诚服，也至少可以被公众所接受。人们曾批评波普尔在这里用了“约定主义”的语言，批评他说接受一个基本语句似乎是一种约定或社会决定；但我想，波普尔思想中听上去像约定主义成分的东西，不过是对那些我们在日常知觉判断中所求

助的隐规范的惯例化本性的承认而已。当一位怀疑论者非要我们"证明"像"我正站在地上"这样的陈述不可时，我们对他的反应的本质，证实了存在着一些要求人们在合适的场合同意这些陈述的社会规范。

维特根斯坦强调说，如果没有这样一些公共规范，没有这样一些为一个团体所共有并构成一种"生活形式"的规范的话，语言，甚至思想本身，都会是不可能的。在维特根斯坦看来，要问我一直在谈论的那种惯例化的证实是不是"确实"正当的，这是荒唐的。在《论确实性》一文中，维特根斯坦指出，对于"猫不生长在树上"这个陈述，哲学家们可以向人提供一百条认识论上的"正当性证明"——但是其中没有一条是从比以下事实更确定(根据"确定"一词的这种惯例化了的理解)的东西出发的：猫不生长在树上。

怀疑论者不仅怀疑过知觉判断，而且怀疑过日常归纳。休谟(我不打算讨论他对"合理的"与"有理的"之间的区分)会说，对于今年冬天美国将下雪(甚至可能下雪)，不存在任何合理的证明(虽然他会补充道，不相信今年冬天美国将下雪是最没有道理的)。但是，我们对一位非要我们"证明"今年冬天美国将下雪不可的怀疑论者的反应，表明存在着这样一些社会规范，它们要求我们就像同意关于站在地上的人以及等臂天平这种日常知觉判断一样，同意上述归纳。

在谈到精密科学中的一些高层次理论时，人们的反应就有点不同了。普通人是无法"证实"狭义相对论的。实际上，普通人眼下甚至不学狭义相对论，或为理解它所必需的(比较基础性的)数学，虽然我们有些大学正开始在一年级物理课中教授狭义相对论。普通人尊重科学家们对这类理论所作的学识渊博的(并为社会所公认的)评价。由于科学理论的可变性，科学家即使对于像狭义相对论这样一个成功的理论，也不可能简单地称之为"真的"。相反，科学家共同体的判断是认为狭义相对论是一个"成功的"——事实

上,像量子电动力学一样,是一个空前成功的——科学理论,它产生了"成功的预见",并"受到了大量实验的支持"。实际上,社会的其他成员都尊重这些判断。这个事例与前面提到的那些惯例化的证实规范的事例相比,区别在于(除了对"真"这个形容词的回避之外),这个事例中存在着专家的特殊作用,和对专家的惯例化的尊重;但这不过是社会中智力劳动分工的一例(且不提各种智力权威关系)。说狭义相对论和量子电动力学是"我们拥有的最成功的物理学理论",这个判断是一些权威人士作出的,而这些权威人士是由社会指定的,他们的权威被一大批习俗和风尚所承认,因而是惯例化的。

最近我想到,维特根斯坦很可能会认为只有能以某种这类"惯例化的"方式加以证实的陈述才可能谈得上是真的(或是对的、正确的、或证明为正当的)。我并不是要提出有什么哲学家曾经主张我们社会中被当作"证明为正当"的任何东西都果真如此。哲学家们一般都区别那些由我们的概念本身所构成的惯例和那些具有某种别的地位的惯例,虽然在如何作这一区别方面还有颇大的争议。我要指出的是,维特根斯坦曾认为,在我们所从事的"语言游戏"中说什么是对的,说什么是不对的,由我们惯例化的证实规范的某个子集所决定。他还认为,在此之外,不存在客观的对或错。虽然这个诠释确实符合维特根斯坦说的不少话——例如,他强调我们如果真要有概念的话就必须"在判断上达成一致"——但我不能担保这是对的。在维特根斯坦对"我们的判断"的说法之中,"我们"是指谁,实在太含糊了;而且我也不知道他的"生活形式"是不是符合我刚才提到的惯例化的规范。但我在读维特根斯坦的《演讲与对话》一书时想到了这个诠释。在这本书里,维特根斯坦既反对心理分析,又反对达尔文的进化论(虽然与实证论者不同,他并没有把这样的语言看成是无意义的,而且他赞赏弗洛伊德的"聪明")。维特根斯坦对心理分析(他称之为一种神话)的态度并不很重要,因为

许多人都认为——在我看来是错误地认为——心理分析或多或少是一种废话。但是他对于进化的拒绝却是令人震惊的。[①]维特根斯坦在达尔文理论和一些物理学理论之间作了不利于前者的对比（“谈到说明，最重要的就是它应当发生作用，它应当使我们能有所预见。物理学与工程技术相联。桥梁万不能断掉”（《美学讲演》，第25页）。他还说，在说服人们时所借助的“根据极为微弱”，“到头来你对有关证实的每一个问题都忘得一干二净，你只知道去确信事情似乎一定是这样的。”

此外，50年代进行的关于“分析性”的大讨论，在我看来同哲学家们的以下愿望有关：他们想为自己的论证找到一种客观的、没有争议的基础。但是为什么哲学家们总倾向于把那么多理应不是“语言规则”或语言规则之结论的东西宣布为是分析的、或“在概念上必然的”，或诸如此类东西呢？我想，答案在于，那种认为存在着一组确定的“语言规则”并且这些规则可以裁定什么是合理的，什么不是合理的观点，哲学家们觉得有两个好处：(1)“语言规则”是基本的惯例化的实践方式（或作为这些实践方式之基础的规范），并因而具有我前面所描述的“公认的”地位；(2)同时，据说只有哲

---

① 维特根斯坦在进化方面的说法是：“人们是根据极为微弱的论据才信以为真的。难道不可能有过如下一种态度——“我不知道。这是一个可能最终得到有力确认的有趣假说”——吗？《美学演讲》，第26页，见西里尔·伯雷特编的《L·W·维特根斯坦：演讲与对话》，伯克利，加州大学出版社，1967年。进化会以什么方式得到“有力确认”，维特根斯坦未置一辞，但他的一段话表明，他想到的是实际看到物种的出现（“有什么人过去看见这个过程发生吗？没有。有什么人现在看见它发生吗？没有。育种的证据不过是杯水车薪。”）。

把维特根斯坦和莫诺两人的态度作一个比较是很有启发的：

达尔文本人表述的选择的进化论要求有孟德尔遗传学的发现，这个发现当然已经作出了。一个理论的内容（或一个观念的内容）是什么意思，这就是一个例证，一个最重要的例证……（一个）好的理论或好的观点的内容即使比它的发明者在当时所知道的也还要广泛得多、丰富得多。完全可以根据这类发展（越来越多的东西投向它的怀抱，即使以前不可能料到它竟会带来那么多东西）来判断这个理论（J·莫诺：“论分子的进化论”，见R·哈勒编《科学革命诸问题：科学中的进步和进步中的障碍》，牛津，1975年）。

学家(而不是语言学家)能够发现这些神秘之物。这个观点在其发生影响的时期,是一个不错的观点,但它一定会四分五裂,而且确实四分五裂了。

凡是认为存在着规定什么可合理地接受、什么不可合理地接受的惯例化规范的观点,我将称之为合理性的唯标准观点。逻辑实证论者,维特根斯坦(至少根据我曾尝试的被认为尚不确定的诠释),以及牛津的"日常语言"哲学家中的有些人(虽然不是全部),[①]都持有一种合理性的唯标准观,即使他们在别的一些问题上(比如是否把不可证实的陈述叫做是"无意义的",有些伦理学命题是不是"概念上必然的")各持己见。

我一开始提到的那个攻势,即拒绝逻辑实证主义的证实原则的那个攻势,之所以是深刻的,恰恰是因为它拒绝为合理性的唯标准观点辩护的每一个企图,是因为它拒绝认为只有可根据标准加以证实的才是可合理地证实的这样一个论题。

---

① 如果有人能提出并论证这样一种观点,认为支配语言行为方式的规范本身是无法用日常的经验研究发现的,他也就能形成一种不承诺对哲学论题进行公开和"唯标准"证实的"日常语言"哲学。斯坦利·卡维尔在《我们必须心口一致吗》一书中朝这个方向走出了重要的一步。他主张为了认识这种规范,可以用一系列他比之于"治疗术"所得到的洞察和现象学所追求的超验知识的"自我认识"。我和卡维尔一样认为,作为一个以英语为母语的人,我关于某些用法是异常还是不异常的知识并不是"外在"的归纳知识——我可以不用证据就知道在我的英语方言中某人是讲"mice"而不是讲"mouses"——但我还是倾向于认为,说话者特有的对自己语言的接近这个事实并不能推广到关于正确性和不正确性的概括。如果我(像卡维尔那样)说,形式为X是任性的这种行话的正确使用规则中包括这样一条:X身上应当有某种"可疑"的东西,那么我是正在提出一种说明我在某些特例方面的直觉的理论,而不是在报告这些直觉。诚然,在心理治疗中也发生这类事情;但在这两种情况下,我都不想赋予自我知识以任何避免他人批判(包括取决于提供对立说明的批判)的性质。如果有人承认这种批判的正当性,那么发现这种规范的活动就开始显得像社会科学或历史学——我论证过,对"科学方法"的传统说明在这些领域中影响甚微。(见我的《意义和道德科学》,劳特利奇和基根·保罗出版公司,1978年。)

不管怎么样,不管语言使用的规范的地位是什么,我看不出有什么理由相信这些规范是"可合理接受的"、"证明为正当的"、"有力确证的"等等的范围的决定因素。

关键在于，虽然我提到的那些哲学家的口气常常显得他们的论证和数学证明或物理学中的判决性实验具有同一种“定局性”似的；虽然逻辑实证论者称自己的工作为科学的逻辑；虽然维特根斯坦对那些“看”不出前维特根斯坦或非维特根斯坦类型的所有哲学活动都毫无意义的哲学家表现出了令人难以置信的傲慢；虽然日常语言哲学家们把别人的每个论证和非日常语言哲学家的论证都称为“洋相”（似乎哲学错误就像算术测验的错误）；但是，没有一个值得一提的哲学立场是可以按我曾描述过的那种一锤定音的、在文化上公认的方式予以证实的。简言之，假如果真只有能按标准得到证实的陈述才是可合理地接受的陈述，那么这个陈述本身并不能按标准得到证实，因而并不能成为可合理地接受的。如果毕竟还有合理性这样的事情的话——我们通过从事说话和论证的活动而全身心地相信某种合理性概念——那么，为如下立场进行辩护，即合理性等同于或恰当地包括于被某文化的惯例化规范判定为该文化之种种表现的东西之中，便是自相反驳的。因为所有这样的论证都不可能仅仅由这些规范确证为正确的。

我决不是认为哲学中不可能有合理的论证和合理的正当性证明，相反，我还不得不承认某种可能对外行人来说显而易见（即使对哲学家不是这样）的事情，即我们不能求助于公认的规范来决定哲学中什么是、什么不是可合理地论证和证明为正当的。现在仍然经常听说的一种观点认为哲学是“概念分析”，而概念本身则决定什么哲学论证是对的，这种观点一旦同认为概念是作为公众语言行为方式之基础的规范或规则的学说相结合，就不过是以下观点的翻版：哲学中一切合理的正当性辩护都是唯标准的，哲学真理（除了“洋相”）是和科学真理一样可公开证明的。从这个问题的整个历史（包括近来的历史）来看，我觉得这样一种观点简直是不合情理的。

适用于哲学论辩的，也适用于关于宗教和世俗意识形态的论

辩。一位聪明的自由主义者与一位聪明的马克思主义者之间的论辩，最终具有和哲学争端同样的性质，即使它涉及更多的经验事实。在宗教、政治或哲学方面我们确实都各有己见，我们都对之进行论证，都对别人的论证提出批判，事实上，即使在科学之中（在精密科学之外），在历史学、社会学和临床心理学之中，我们也进行具有完全同样性质的论辩。诚然，逻辑实证论者扩展了他们对"科学方法"的描述，把这些领域也包括进来；但这范围扩展得这样广，已经无法表明有什么东西是显然不属于其中的（见第八章）。

我将提到，实证论者曾让步说，证实原则是"无认知意义的"。他们说，这个原则是一个建议，因而是没有真假的。但是，他们论证自己的建议，而这些论证则是（而且只能是）一些旁观者[①]。所以问题仍然存在。

总而言之，逻辑实证论者和维特根斯坦（或许还包括后期蒯因）所做的，是创造出一些不容纳合理的哲学活动的哲学。正因如此，这些观点是自相反驳的；而且，我一直在讨论的那个小小攻势，代表了那种哲学家们称为"先验论证"的重要论证：对合理性的本

[①] 在捍卫作为一种建议的证实原则时所提出的最弱的论证是说它"阐明了""前分析的"有意义性概念。（要了解关于这一主张的讨论，见我的"怎样不去谈论意义"一文，载拙著《心灵、语言和实在——哲学论文集第二卷》，剑桥大学出版社，1975年。）赖辛巴赫捍卫以下形式的证实原则（在《经验与预见》一书中），它保留同行为有关的所有意义差异。在反驳一个明显的异议（它认为对神祇的非经验信念——赖辛巴赫用埃及人的猫类崇拜作为例子——能够改变行为）时，赖辛巴赫提议将"猫是神圣动物"转译为"猫激起猫类崇拜者的敬畏感情"。显然，接受这样一个替代句子，在一个猫类崇拜者那里，并非对行为毫无改变！

最有意思的观点要数卡尔纳普的了。在卡尔纳普看来，所有合理重建都是建议。仅有的事实问题，是同接受这个或那个合理重建的逻辑的或经验的结论有关的。（卡尔纳普把合理重建比作为飞机选择一台发动机。）他得出的结论是在哲学中，人们应当容忍不同的合理重建。但是，卡尔纳普所说的这个"容忍"原则，预设了证实原则。因为，说所有合理重建都不是唯一正确或唯一符合"事情""实际存在"的方式的观点，说一切"外部问题"都没有认知意义的观点，正是"证实原则"。把"容忍原则"运用于"证实原则"本身，将是循环论证。

质进行论证(典型的哲学家的任务)是这样一种活动,它所预设的合理的正当性辩护的观念要宽于实证主义的观念,甚至宽于惯例化的唯标准的合理性。

无政府主义是自相反驳的

现在我们来讨论一种非常不同的哲学倾向。托马斯·库恩的《科学革命的结构》(简称《结构》)迷住了大量读者,也吓坏了大部分科学哲学家,因为它强调的东西看上去是接受科学理论过程中的非理性的决定因素,因为它使用像"皈依"和"格式塔转换"这样一些术语。事实上,库恩在科学理论和应当如何看待科学活动方面提出了许多重要观点。在别的地方我已经表示过,我相信"范式"、"常规科学"和"科学革命"这些概念的重要性;在这里,我想集中谈我觉得同库恩的观点不同的地方,集中于我在别处所称的"库恩的极端相对主义"。

令那些冒失的读者们陶醉不已的是按以下理解阅读库恩著作,以为他正在说科学中没有合理的正当性证明之类的事情,而只有格式塔转换和皈依。库恩曾拒绝对《结构》一书的这种诠释,后来还引入了一个"非范式的合理性"的概念,这个概念与我所谓的"非唯标准的合理性",如果不是一样的话,也是密切相关的。

大多数读者以为自己在《结构》一书中发现的那种倾向,在保罗·法伊尔阿本德《反对方法》一书中确实表现出来了。法伊尔阿本德像库恩一样强调不同的文化、不同的历史时代是如何产生出不同的合理性范式的。他提出我们的科学合理性概念的决定因素大部分是我们将称之为非理性的东西。实际上,(虽然他没有这样明说)他提出现代科学技术的合理性概念是它受自己的标准欺骗的产物。(我认为在米歇尔·福科的书中我也发现了一种类似的语气。)他比库恩和福科走得更远,提出即使对我们科学的自负的工

具性的优越感，也是某种骗局。法伊尔阿本德声称，要解除你的病痛，信仰治疗师比医生更为有效。

我想谈论的并不是这些惊世骇俗的激进之论，虽然正是因为这些论调法伊尔阿本德才称他的立场为"无政府主义"的。我想要讨论的是库恩在《结构》一书和后来的论文中都提出的一个论点，也是法伊尔阿本德在《反对方法》一书和一些技术性论文中都提出的一个观点，这就是"不可公度性"的命题。我想要说的是，这个命题，和逻辑实证主义关于意义和证实的命题一样，也是一个自相反驳的命题。简言之，20 世纪这两种最有影响的科学哲学，当然也是普遍地引起科学家和非哲学家兴趣的两种科学哲学、有教养的一般读者可能耳有所闻的仅有的两种科学哲学，都是自相反驳的。当然，作为一名科学哲学家，我觉得情形居然如此有点麻烦。过不久我们将讨论是什么造成了这种局面。

不可公度性命题认为，在另一个文化中使用的术语，比方说 17 世纪科学家所用的"温度"这个术语，在意义或指称上不能与我们今天拥有的术语和表达式划上等号。比如库恩指出，拥有不同范式的科学家居住在"不同的世界"。1900 年前后使用的"电子"指称的是这个"世界"中的对象；而今天使用的"电子"则指称完全不同的另一个"世界"中的对象。这个命题被认为不但要运用于所谓"理论语言"，而且要运用于观察语言；实际上，在法伊尔阿本德看来，日常语言干脆是一种假的理论。

这回对这个命题的驳斥是说，假如这个命题确实是真的，那么我们根本就无法翻译别的语言——甚至以前阶段的我们的语言。而如果我们对有机体的声音全然无法诠释，那么我们就没有根据把他们当作思想家、说话者甚至人。简言之，如果法伊尔阿本德（以及库恩——在其大力主张"不可公度性"的时候）是对的，那么其他文化的成员，包括 17 世纪的科学家，就可能被我们在概念上描述为仅仅是对刺激（包括那些竟然不可思议地同英语或意大利语相

似的声音)作出反应的动物。先对我们说伽利略具有"不可公度的"概念,然后进一步去详细描述这些概念,这完全是自相矛盾的。

斯马特在一篇同情法伊尔阿本德观点的论文中也提出了这个问题:[①]

> 为了看见水星,我们必须把望远镜指向比方说那棵树顶端的上空,而不是如牛顿理论预见的那样,指向那只烟囱罩顶端的上空,这肯定是一个中性的事实。对树木、烟囱罩和望远镜,人们肯定也能以一种同选择牛顿理论还是选择爱因斯坦理论无关的方式进行谈论。但是,法伊尔阿本德很可能退而承认,为建立我们的望远镜理论,我们使用的是欧几里得几何和非相对性光学。他会说,这并不是关于我们的望远镜、树木和烟囱罩的实在的真理。然而,为了讨论广义相对论的观察检验而以上述方式考虑是正当的。因为我们基于理论上的根据知道,即使我们利用这种计算上的便利,我们的预见也将不受影响(只要没有超过观察误差的限度)。

但是斯马特营救之举的麻烦在于,我即使说这些"预见"是相同的,也必须理解一些欧氏非相对性语言。如果每个词都有不同意义,一个预见能在什么意义上"不受影响"呢?如果我无法找到一本将17世纪意大利语和现代英语挂起钩来、不仅在这种17世纪语言本身之内、而且在它的语言外背景中对它获得系统理解的翻译手册,那么,即使对于17世纪意大利语中的一些逻辑助词(用以表达"如果——那么","并非"等等的单词),或诸如此类的东西,我又

---

① J·J·C·斯马特:"关于说明的种种对立观点",收于R·柯亨和M·瓦托夫斯基主编的《波士顿科学哲学丛书·第2卷:纪念菲利普·弗兰克》一书(纽约人文科学出版公司,1965年)。

怎么可能翻译呢?即使我是一个能运用这两种理论的说话者(如斯马特设想的那样),我又怎么能有根据地把牛顿理论中任何词项同我的广义相对论中的任何词项等同起来呢?

当我们把蒯因和戴维森关于意义和翻译活动的一些说法运用于我正在讨论的观点时,这个观点就成了更加激烈的争论的焦点。一旦人们退一步承认我们可以找到一种翻译方法,它在17世纪文本那里,至少在由我们的兴趣和这翻译的用途所确定的上下文中,最"行之有效"的,那么,在这种上下文中说这翻译并未"真地"把握住原作的含义或指称,有什么意义呢?除了我们的翻译方法和我们对这些方法在经验上的适当性的或明或暗的要求之外,我们居然还必定或可能拥有含义或指称之相同性的标准——这怎么可能呢?对于说某译本未能完全把握原作含义或指称这样一个论断,人们可以把它理解成对如下情况的承认,即有可能找到一种好一些的译本;但是,说所有可能的译本都无法把握"实在的"含义或指称,这话听起来有意义,实际上只是一种错觉。我们在为了诠释而把不同的表达式等同起来的时候,要用到一种关系,或一个关系家族(这两者都有点含糊),而同义性就是仅仅作为这样一种关系或关系家族而存在的。那种认为在一切可行的互译惯例之外还存在"实在的"同义性之类的东西的看法,已经被人视如神话,予以抛弃了。

假定有人对我们说,德语单词"Rad"不可译作"轮子"。如果他接着说他的译本不够完善,我们自然而然地会等待他指出这译本可以怎样用注释之类的东西来加以改进和补充。但是,如果他接着说"Rad"可以译作"轮子",然而它实际上并不指称轮子,甚至不指称在你的概念系统中可辨认的任何对象,从这说法中我们得到了什么呢?说一个单词A可以译作"轮子",或诸如此类的东西,就是说,只要该翻译可以信赖,A就指称轮子。

不可公度性命题之所以如此蛊惑人心的原因,除了所有不融

贯的观点好像都有的那种魅力之外，可能在于把概念和观念混为一谈的倾向。只要分析的/综合的这个区别是模糊的，概念和观念的两者间区别就也是模糊的；但是所有诠释都包含这样一种区别，即使它是相对于诠释本身的。在我们翻译一个像温度这样的词项时，我们把指称（只要我们坚持自己的翻译，那么也把被翻译的表达式的含义）等同于我们自己的术语"温度"，至少当我们在那个上下文中使用它的时候。（当然，我们可以用各种措施，比如专用注释，去限定和澄清我们在那个上下文中使用"温度"或不管什么词项的方法。）在此意义上，我们把这个"概念"同我们自己的温度'概念"等同起来了。但是这样做，是同以下事实并行不悖的，即17世纪的科学家，或随便什么人，可能有过一种与我们不同的有关温度的观念，也就是一套与我们不同的有关温度的信念，不同的"知识意象"，还有，不同的有关许多其他物质的基本信念。这些观念的不同，并不证明（像有时人们所主张的那样）不可能"确实正确地"翻译任何术语；相反，如果我们不能进行翻译的话，我们就说不出观念是不同的，是如何不同的。

但是，有人会问，如果观念总是被证明为各不相同的，我们怎么还能知道一种翻译方法是"行之有效"的呢？从维科至今，已经有各种思想家提出过对此问题的回答，认为诠释的成功，并不要求被译者的信念被认为是与我们的信念相同的，但它确实要求这些信念被认为是我们能够理解的。各种关于诠释上的宽厚或"疑惑的好处"的格言，诸如"在诠释时，使他们被看作是相信真、热爱善的人"，或"在诠释时，使他们的信念根据他们所受到的教育、所具有的经历而被认为是合理的"，或维科本人的尽量体现被诠释之人的人性的要求——所有这些格言的基础，都在这里。这个世界上的各种文化一方面各自经历或慢或快的变化，另一方面在历史上发生相互作用；而有关这样一个世界上的人类经验一个基本事实便是，作为一种普遍的人类经验，我们是能够做到这一点的；是能够把某

个他人的信念、欲望和表达诠释得确有某种意味的。

无怪乎库恩和法伊尔阿本德拒绝任何有关科学知识中的会聚的想法。既然我们同以前的科学家们谈论的不是同样的东西，我们也就并非在对同样的微观对象或宏观对象得到越来越多的知识。库恩强调说科学仅在工具方面取得“进步”；我们越来越善于把人们从一处送往另一处，等等。但这个说法也是不融贯的。除非像“把人们从一处运往另一处”这样的特殊说法的指称保持某种程度的确定性，否则，我们怎么能以任何稳定的方式理解工具性成功这一概念呢？

我刚才运用的论据本质上是与库恩关于经验知识的先决条件的著名论据有关的。在回答关于未来可能毫无规律、可能驳倒我们已作的每一个“归纳”的论点时，库恩指出，如果毕竟还存在着未来的话——至少，如果存在着对我们而言的未来、我们作为思想者能够把握、能够用概念进行处理并进而说我们的预见是真还是假的未来的话——那么，其实人们原先就不得违反许多规则性。要不然我们为什么要称之为未来呢？例如，当我们想象珠子以某种“不规则”方式从一只瓮里取出时，我们忘记了，如果不依据许多规则性的话，我们就连它们是珠子、它们被取出的顺序也说不出来。要进行比较便预设了某种可公度性的存在。

为回答上述所有批评，库恩和法伊尔阿本德还可以采取一个步骤，但这个步骤他们不会乐意采取，那将意味着引进某种观察/理论二分法。他们可以在观察事实方面姑且承认可公度性、可译性，甚至会聚性，而把不可公度性命题限于理论性词汇。即使那样，也有问题（为什么我们不按照拉姆齐的方式用理论术语同观察词汇的关系来描述理论术语的意义呢？）。但库恩和法伊尔阿本德理所当然地拒绝这个可选择方案，因为，诠释上宽厚的原则不仅在“理论语言”中，而且在“观察语言”中，都是普遍需要的。比方说，请考虑“草”这个常用词，不同的说话者，依其生活的地点和时间，对

草具有不同的知觉样本(不同的地方,草有不同的颜色和不同的形状)和观念。虽然所有说话者都一定知道草是一种植物,否则会被说成具有完全不同的概念,但今天植物这个观念中包括光合作用,而200年前的植物观念则不包括光合作用。如果没有诠释上的宽容引导我们把200年前的"植物"等同于今天的"植物"(至少在日常的语境中),把200年前的"草"等同于今天的"草",对200年前这个词的指称就会根本无话可说。不仅自然物种的词项在翻译时依赖于宽容原则,"面包"这个人造物词项也提出完全同样的问题。要没有诠释上宽容,就连像"红色"这样一个简单的颜色词,我们也无法在不同的说话者之间按同样意义使用。对于话语,我们总是作为一个整体来诠释的;就像"理论"术语的诠释依赖"观察"术语的诠释一样,"观察"术语的诠释也依赖于"理论"术语的诠释。

我上面提供的,再说一遍,是一种超验的论证。我们一些根本性的观念,使我们不能不把我们过去的自我、我们的祖先以及其他文化中过去和现在的成员,而不仅仅我们当前的这个片段,当作人来对待;而这意味着——正如我强调过的——赋予他们以共有的指称和共有的概念,而不管我们同时赋予他们的观念是如何不同。在诠释活动成功进行的范围内,我们同他人不仅拥有共同的对象和概念,而且拥有对合理的东西、自然的东西等等的共同的观念。因为,别忘了,对一个诠释方法的全部正当性证明,就是它使得他人的行为在我们看来至少有低限度的合理性。不管我们心目中知识的意象、我们对合理性的观点多么不同,即使是最希奇古怪的文化,只要我们毕竟还能成功地予以诠释的话,在什么是合理的这个问题上,我们也与之拥有大量共同的假定和信念。

为什么相对主义是自相矛盾的

在哲学家中间,(彻底的)相对主义是自相矛盾的这一点,是不

言而喻的。不管怎样，主张一种观点，但同时又主张没有任何观点比别的观点更有根据、更加正当，这难道还不是自相矛盾的吗？阿兰·加芬克尔曾相当机智地表达过这个观点。一次，在同他的加州学生谈话时，他模仿他们的习惯用语说："你们的出发点可能与我不同，但我知道相对主义对我来说不是真的"……如果任何观点都与任何别的观点一样好，那么，相对主义是假的这个观点为什么就不和其他观点一样好呢？

相对主义理论在目前的大量兜售（其卖主还是一些才智颇高的思想家）表明，上述简单的反驳是不够的。为什么一位明智的相对主义者要承认每个观点都和任何其他观点一样（对他来说）是真的呢？他无法阻止你（或阿兰·加芬克尔）说他的观点对你来说不是真的（或对你来说是正当性未被证明的，或诸如此类的东西），但是如果他老谋深算的话，他可以反驳说，对你来说是真的东西，（在他看来）远远不如对他来说是真的东西那么显而易见。不管怎样，有什么概念能比人们自己的概念来得更为显而易见呢？在这种情况下，把真的、证明为正当的等等看作是相对的概念，难道还真是自相矛盾的吗？

回答是，它确实是自相矛盾的，但它也确实要求提供一种比加芬克尔提供的单线条的（但也是非常巧妙的）论证更精巧的论证。需注意的要点是，如果一切都是相对的，那么相对的也是相对的。但这需要作不少说明！

柏拉图可能是第一个运用我想到的那种论证的人（用来反驳普罗泰戈拉）。普罗泰戈拉（他显然是一位顽固不化的相对主义者）认为，当我说 X 时，我其实应该说"我认为 X"。这样，当我说"雪是白的"时，普罗泰戈拉会说，我的意思其实是希拉里·普特南认为雪是白的，而同样的话在罗伯特·诺齐克那里则意为罗伯特·诺齐克认为雪是白的。这同样的观点用更为精致的说法来表达可能是这样的，即当我说"雪是白的"时，我是在用这句话表示雪

是白的对我而言是真的，而当罗伯特·诺齐克说同样的话时，他通常是在表示，雪是白的对他而言是真的（或至少只有在这句话被证明为对他而言是真的时候，他才会认为它曾是正确的）。这意味着（根据普罗泰戈拉的观点），没有一句话对我和对别人是有同样意义的；我们前面看到，在相对主义和不可公度性之间，存在着密切的联系。柏拉图的反论证是这样的：如果每一个X陈述的意义是"我认为X"，那么，（根据普罗泰戈拉的观点）我其实应该说

（1）我认为我认为雪是白的。

但是，添加"我认为"的过程是可以永远重复的！这样，根据普罗泰戈拉的观点，"雪是白的"最终的意义不是（1），而是

（2）我认为我认为我认为我……（"我认为"的数量为无穷多）雪是白的。

柏拉图把这个论证当作一种归谬论证。但柏拉图的这个论证并不像初看之下那么完善。为什么普罗泰戈拉不可以同意说，他的分析适用于它本身呢？这并不意味着它一定要无限次地运用于自身，而只意味着它可以无论多少有限次地运用于自身。但柏拉图看到了某些非常深刻的东西。

人们第一次碰到相对主义时，该观点看上去非常简单。根据其自然的最初的表述，这个观点认为每一个人（或，在现代的"社会学的"表述中，每一个文化，或有时是每一段"话语"）有其自己的观点、标准、预设，而真理（以及正当性证明）是相对于这些东西而言的。当然，X相对于这些东西是否为真的（或其正当性得到辩护的）这一点本身是某种"绝对的"东西，对此人们认为是不言而喻的。

按照像福科那样的现代结构主义者的观点，相对于一段话语

的正当性证明本身似乎是相当绝对的——也就是说，根本不是相对的。但是，如果形式为“X相对于某人P是真的或证明为正当的”这样的陈述本身是绝对地真或者假的，那么，毕竟存在着一个绝对的真理（或证明为正当的）概念，而不仅仅是为我之真，为诺齐克之真，为你之真等等。但一个彻底的相对主义者会不得不说，X相对于P是否真这本身也是相对的。在这里，正如柏拉图指出的，我们对于这观点的意图的理解，也开始动摇了。

柏拉图用以抨击相对主义的思路，真到最近才好像有人仿效。但是维特根斯坦——尤其在其私人语言论证中（第三章曾提到）——曾对此作了出色的发挥。

大部分评论家都把私人语言论证仅仅读作一种反对真理的“摹本论”的论证。维特根斯坦对于以下一点的出色证明，即指称的相似论即使在对感觉的指称的问题上也是无效的，当然确实是对于形而上学实在论的持久抨击的一个部分。但我更愿意把这个论证读成这样一对相当传统的论证（至少康德对这两者都会予以赞同的），它们反对两种观点，一个是实在论的观点，另一个是相对主义的观点；因为，企图把这整个论证理解成一种反实在论的论证，会使它看上去非常费解。

维特根斯坦想要抨击的那种相对主义，人们称为“方法论的唯我论”。“方法论的唯我论者”是一位这样的非实在论者或“证实主义者”，他同意把真理理解为是在某种程度上同合理的可接受性相联系的，但他又认为所有的正当性证明最终都根据我们人人对之具有私人知识的那些经验。这样，我自己知道我的哪些经验会证实雪是白的，诺齐克自己知道他的哪些经验会证实雪是白的；每一个陈述对于每一个思想者来说，都具有不同的意义。

在我看来，维特根斯坦的论证是一个反对整个相对主义的出色论证。这段论证是说，相对主义者到头来都无法理解是对的和认为他是对的这两者之间的区别；而这意味着，断言或认为与发声

（或产生心理意象）这两方面之间的区别，最终是不存在的。但是这意味着（根据这个观点）我根本不是一个思想者，而仅仅是一个动物。主张这样一种观点无异于某种精神自杀。

为了看出维特根斯坦是对的，我们来考虑一下（维特根斯坦并没有这样做），相对主义者会如何设法作出维特根斯坦认为他们作不出的那种区分，也就是"对的"和"认为他是对的"这两者之间的区分。

相对主义者会借用真理是合理的可接受性的一种理想化这一观点。他会主张，如果在我足够仔细地观察、足够多时间地推理或诸如此类的情况下，"X 对我来说是被证明为正当的"可能是真的，那么，X 就是对我来说是真的。但是，形式为"如果我将……，那么我会认为如此这般"的假设条件句，像所有陈述一样，在不同哲学家那里得到不同的诠释。

一个形而上学实在论者会把关于"如果，则如何"的陈述当作本身是绝对地真或假的，而不依赖于我们对这些陈述的接受或拒绝是否将得到辩护。如果相对主义者以这种实在论的方式诠释有关如此这般条件下他会相信什么的陈述的话，那么他已经承认了某一类绝对真理，从而也拒绝了成为一个相对主义者。

非实在论者或"内在的"实在论者把条件陈述看作（像所有其他陈述一样）是我们主要通过把握其正当性证明条件来加以理解的陈述。这并不意味"内在的"实在论者放弃真与证明为正当之间的区别，而意味着我们对真（理想化的证明为正当）的把握如同对任何其他概念的把握一样，是通过对某些这样的因素的理解（多半是隐含的）而实现的，这些因素使得说某事为真是可被合理地接受的。相对主义者对于有关他在理想条件下会相信什么的陈述，能够以这种非实在论者的或"内在的"实在论的方式予以诠释吗？

我们还记得，非实在论的观点，正如我描述过的（在第三章），采纳一种客观的合理可接受性概念。非实在论者反对说真理是对

于一个"现成世界"的符合的观点。这便是他之所以是一个非(形而上学)实在论者的原因。但是,反对形而上学的真的"符合"论与把真或合理的可接受性当作主观的,根本不是一回事。N·古德曼——他把真与合理的可接受性看作是既可以运用于陈述句也可以运用于艺术作品的一个更一般的谓词"对"的两个种概念——曾简洁明了地提出过这个观点:

> 因此,简单地说,陈述的真和描述、表征、例证、表达的对——设计、绘画、遣词用句、韵律节奏的对——首先是一件有关适合的事情:以这种或那种方式适合于所指称的东西,适合于其他再现方式,或适合于组织的类型和举止。使一个文本适合于一个世界,使一个世界适合于一个文本,以及使一个文本自我适合,或使之适合于其他文本——一旦承认了文本在创造它们所适合的世界中的作用,这几个方面的区别便渐渐消失了。而了解或理解,则被看作不是仅仅对真的信念的获得,而包括对各种各样适合的发现与发明。

然而,相对主义的全部目的,它的极为明确的特征,正在于否认任何可理解的客观的"适合"概念的存在。这样,相对主义者便无法理解用客观的证明为正当之条件来进行的有关真理的讨论。

在这种情况下,用条件句来阐明是对的和认为某人是对的这两者间区别的企图,便没能成功,因为正像对任何其他陈述一样,对这些条件句,相对主义者也没有任何客观的对的概念可用。

最后,如果今天的相对主义者像古代的普罗泰戈拉一样干脆决定把打落的牙齿往肚里吞,说在"我是对的"和"我认为我是对的"之间不存在任何区别——说即使在某人自己那里也无法区别被证明为正当和认为某人被证明为正当——那么,根据这样一种观点,除了发出声音以求自己感到是对的之外,还说些什么呢?除

了在心里产生意象和类似句子的东西以求主观上感到是对的之外，还想些什么呢？到头来，相对主义者不得不否认任何思想是有关任何东西的，不管在实在论的意义上还是非实在论的意义上；因为，在心想某人的思想是有关某物的，与实际上心想该物之间，他无法作出区分。简言之，相对主义者未能看到的是，某种客观的“对”的存在，是思想的一个前提。

在我刚才分析的那个论证（柏拉图 维特根斯坦）和我归之于蒯因与戴维森的那个反对不可公度性的论证之间，存在着一种颇有意思的联系：蒯因和戴维森所论证的实际上是，一个前后一致的相对主义者根本不应该把别人当作说话者（或思想者）（如果他们的“声音”是那个“不可公度的”，那么它们就仅仅是“声音”），而柏拉图和维特根斯坦所论证的则实际上是，一个前后一致的相对主义者不能把他自己当作一个说话者或思想者。

何 以 至 此？

我刚才放在你们面前的那些论证使我确信，本世纪产生的两种名声最广的科学哲学都是自相矛盾的。（当然，两者都不仅仅是一种“科学哲学”。）这自然而然地引起我反省这种状况的意义。这些观点是怎样产生的呢？

我曾回顾说，逻辑实证主义与在它之前的马赫主义既相连续，又相区别。马赫的实证主义或“经验批判主义”实际上多半是用另一种行话对休谟经验主义的重复。马赫的出色才华，他的固执而热情的风格，以及他的科学声望，使得他的实证主义成了一大文化争端（列宁担心布尔什维克会倒向“经验批判主义”，写了一本与之论战的书）。爱因斯坦—— 他对狭义相对论的诠释实际上是操作主义的（与他对广义相对论的诠释形成显著对照）——承认他对于同时性概念的批判多亏了休谟和马赫，尽管令人失望的是，马赫完全拒

绝狭义相对论。

然而导致逻辑实证主义出现的最惊人的事件是演绎逻辑中发生的革命。在1879年，弗雷格发现一种包括今天的标准“二阶逻辑”的算法，一种机械的证明程序。这个程序对于初等演绎理论（“一阶逻辑”）来说，是完备的。人们居然能够写下一种用于证明一阶逻辑的全部有效公式的算法——一种不要求对整个人类的心理作重要分析与模仿的算法——这是一个了不起的事实。它激起了这样一种希望，即，对所谓“归纳逻辑”人们也可以做同样的事情——“科学方法”也可以被证明为是一种算法，而这两种算法——用于演绎逻辑的算法（当然，推广到高阶逻辑时，它就显得不完备了）和用于归纳逻辑的尚待发现的算法——不仅对科学的合理性，而且对值得一提的所有合理性，都将无所不包地作出描述或予以“合理重建”。

在我刚刚开始我在普林斯顿大学的教学生涯时，我得以结识鲁道夫·卡尔纳普，他那时在高级研究院呆了两年。在一个难忘的下午，卡尔纳普向我描绘他是怎么成为一名哲学家的。卡尔纳普向我解释说，他曾是一名物理学研究生，在弗雷格的讨论班上学习逻辑，教材是《数学原理》。（试想，居然在弗雷格那里学罗素和怀特海的《原理》！）卡尔纳普迷上了符号逻辑，同样也迷上了狭义相对论。所以他决定以《原理》的符号对狭义相对论进行形式化，并以此作为他的论文。卡尔纳普对我说，只是因为耶拿的物理系不接受这篇论文，他才成了一名哲学家。

今天，大量否定性的成果，包括N·古德曼的一些颇为严格的考察，已经告诉人们不可能存在一种完全形式的归纳逻辑。归纳逻辑的有些方面是可以形式化的（虽然这种形式化的恰当性是有争议的），但是永远存在着对“有理性”作出判断的必要，而不管这些判断是通过对词汇的选择（或更确切些说，把词汇分为“可预测的”谓词和“不可预测的”谓词）而确立的，还是以别的什么方式确立

的。今天，实际上没有任何人相信，存在着一种纯形式的科学方法（关于这个问题，请见第八章）。

卡尔纳普告诉我的故事为以下想法提供了佐证，即起着关键作用的正是形式化在演绎逻辑这个特例中的成功。如果这个成功激发了逻辑实证论的兴起的话，那么，对归纳逻辑的形式化的失败，对不存在用于经验科学的算法这个事实的发现，难道不曾激发过“无政府主义”的兴起么？

我不想强调这个主张；不管怎么样，其他因素可能也起着作用。在库恩越来越钝化他的观点的同时，法伊尔阿本德和米歇尔·福科两人却倾向于把它推向极端。他们的想法中有某种政治因素：法伊尔阿本德和福科都把我们目前的惯例化的合理性标准同资本主义、剥削甚至性压抑挂起钩来。极端的相对主义在今天之所以对人们具有魅力，显然有许多不同的理由，而认为所有现成的惯例和传统都是坏的这种观点，便是其中之一。

另一个理由是某种科学主义。逻辑实证论的科学主义是相当大言不惭的；但我认为，在相对主义背后，也隐藏着一种科学主义。虽然从来没有什么“无政府主义的”思想家完全接受过以下理论，认为对合理性来说，所存在的不外乎你的本地文化所认为存在的东西，但这个理论是他们倾向的自然极限：这是一种还原论的理论。认为界定合理性的是一种理想的计算机程序的观点，是一种被精密科学激发起来的科学主义理论；认为界定合理性的仅仅是本地的文化规范的观点，则是一种被人类学激发起来的科学主义理论。

我这里不打算讨论乔姆斯基语言学在某些人那里所引起的那种期望，以为认知心理学将发现界定合理性的内部算法。我本人认为，这种思想风尚像逻辑主义对符号化的归纳逻辑的希望那样，也将遭受失望。

以上所说的一切都表明，当今哲学的问题之一，便是从19世

纪继承而来的科学主义——这是一个不止影响一个思想领域的问题。我并不否认逻辑的重要性，不否认形式化研究在确证论和自然语言的语义学等等方面的重要性。但我确实倾向于认为，它们对于哲学来说是相当肤浅的，而只要我们在形式化的牢牢支配之中，徘徊于我曾描述过的那两种科学主义之间便是可想而知的事情。这两种科学主义都企图回避这个问题：对理性的范围提供一个清醒而人道的描述。

第六章　事实和价值

从最广泛的意义上来看，事实和价值这一课题与我们每个人都息息相关。就此而论，它与许多哲学问题大相径庭。例如，对确实是有一个实在的世界，还是只不过看上去有一个世界这类问题，大多数受过教育的男女都不会以为他们非回答不可。从多数人的生活角度来看，语言哲学，认识论，甚至形而上学中的一些问题，无论多么有趣，都是有点可有可无的。但是，事实和价值的问题则是一个必须选择的问题。每个有头脑的人都不得不对之持有一个实实在在的观点，这种观点与他们想象的观点或许相同，也可能并不相同。如果事实和价值的问题对有头脑的人们是一个必须选择的问题，那么，对这一问题的一个特定的答案，即，事实和价值是两个完全不同的领域，事实陈述和价值判断非此即彼的二分法是绝对的，这种答案就已具有了某种文化惯例的地位。

我之所以称这种二分法为一种文化惯例，是想指出，如果这个公认答案在相当长时间内继续成为公认答案，而不管哲学家们对此会说些什么，也不管这种公认答案对不对，这将是一件不幸的事。即使我有办法使你相信，事实-价值的二分法是没有公理根据的，使你相信，这是一个无法予以合理辩护的二分法，或者，即使某些比我出色的哲学家能够通过一个绝对权威性的论证（在哲学中当然不会有这类论证）来表明这一点，但是当你下次出门上街，出席鸡尾酒会，或者到你偶然加入的某个审议机关作讨论发言时，你还是会发现某个人在对着你说："那应该是一个事实陈述还是一个

价值判断?”认为在某些事情是好还是坏，是较好还是较坏等等的问题上不存在既成事实的观点，从某种意义上讲，已经被惯例化了。

我的论证策略不会是一个新的策略。我打算在关于事实和价值的争论中，为一个多少有点失去信誉的主张恢复名声。这个主张的内容是强调，事实和价值的二分法至少是极为模糊的，因为事实陈述本身，以及我们赖以决定什么是、什么不是一个事实的科学探究惯例，就已经预设了种种价值。

之所以说上述论证是一个多少有点失去信誉的主张，是因为它受到了一个明显的反驳，对于主张科学预设了价值的观点这个反驳是一个防御性的让步。事实-价值二分法的辩护者退而承认，科学的确包含有价值，例如，科学预设着我们需要真理。但是他们辩解说，这些价值不是伦理价值。我将设想这样一位假想敌，他认为科学只预设一种价值，即真值本身。

如我们所知，真理不是一个简单的概念，那种以为真理是某个(独立于心灵和话语的)“真正”“存在”物的被动摹本的观念，即使仍旧深深地占据着普通人的头脑，但经过康德、维特根斯坦和其他哲学家们的批判，已经土崩瓦解了。

有些哲学家为了论证在真理方面并没有真正的哲学问题而诉诸于等价原则。根据这个原则，说一个陈述是真的，等价于断定该陈述。另有一些人诉诸逻辑学家塔斯基的工作。他表明，在给定一种形式化的语言(运用符号逻辑表达某些陈述的一种形式记号)的情况下可以怎样用一个更强的语言(一种所谓“元语言”)来为该语言定义“真”。[1]

塔斯基的工作本身是建立在等价原则上的：事实上，他关于一个成功的“真”的定义的标准，是为一切形式为“P”是真的当且仅

[1] 对塔斯基的非技术化的解释，见拙著《意义和道德科学》，第一编，第一讲。

当 P 的语句，例如，

(T)"雪是白的"是真的，当且仅当雪是白的

提供一些元语言的定理(其中 P 是该形式记号的一个语句)。

但是等价原则在哲学上是中性的。塔斯基的工作也是如此。根据任何真理论，"雪是白的"等值于"'雪是白的'是真的"。

实证主义哲学家也许会回答，如果你懂得上述(T)，你就懂得"'雪是白的'是真的"意味着雪是白的。如果你不理解"雪"和"白"，他们也许会补充说，你才确实出了毛病！但是问题不在于我们不懂"雪是白的"，而在于，我们不懂究竟何为懂得"雪是白的"。这是一个哲学问题，(T)等式于此未置一词。

说实在的，这难道不符合我们对这类事情的直觉吗？如果某人眼神狡黠地走到我们跟前说："你们想知道'真理'吗？"我们的反应一般是对他满腹狐疑。而我们之所以这样(除了他的眼神之外)，恰恰是因为只要我们对此人所持的合理的可接受性标准心中无数，只要我们对以下问题一无所知，他那番想使我们知道真理的表白就会使我们一无所获。这些问题是：他们所认为的进行探究的合理方式是什么，他们的客观性标准是什么，他们认为何时终止一个探究是合理的，他们认为当他们要对自己感兴趣的不管什么问题接受这个或那个结论时，什么样的根据可以为之提供一个好的理由。运用到科学中，我想说，告诉我们科学"在追求发现真理"其实只是一个纯形式陈述。它只相当于说，如果雪不是白的，科学家就不想断定雪是白的，相当于说，如果电流没有通过电线，他们就不想断定有电流通过电线，等等。将确定雪是否白的科学方法与确定雪是否白的其他方法区分开来，将确定电流是否通过电线的科学方法与确定电流是否通过电线的其他方法区分开来，需要一些合理可接受性的标准。只要我们对这个合理可接受性的标准系统一无所

知，这些纯形式的陈述也将是完全空洞无物的。

如果那种将我们的信仰体系和非概念化的现实相比较，以便发现它们是否相配的想法是无意义的，那么科学追求发现真理的论断仅仅意味着，科学试图构造这样一幅世界图景，它能最为理想地满足合理的可接受性的某些标准。说科学试图构造一个真的世界图景这个陈述本身是一个真陈述；科学的目标只有通过隐含在科学之内的合理的可接受性标准才能获得实质内容。简言之，对于前面所列举的假想的立场（这个立场，除了正确地指出科学有其他附带的目标之外，还认为科学的唯一目标是发现真理），我的回答是：真理不是终极之物，真理本身还要从我们的合理可接受性标准那里获得生命，而如果我们希望发现真正隐含在科学之中的价值，这些标准是务必注意的。

为说明问题，我来设想一个观点分歧的极端例子。我要设想的不是普通的科学上的分歧，尽管我希望我们对这种分歧的反应能使我们得以发现科学价值的本质。

在我要描述的论述中，这个分歧是在第一章中所讨论的假定——我们全都是钵中之脑——上发生的。我们已经论证过，这个假设不可能是真的；但是我们将假定，这个论证并没有使这个分歧中的某一方信服（这并不是不可能的，因为哲学论证从未使任何人信服过）。简言之，这个假定是说，像我们前面描述过的那样，一切都是一种集体性的幻觉。

当然，如果一切全都按这种方式成为同一种集体幻觉的话，有许多人会对此抱无所谓的态度。例如，对情人们来说，这一切便可能与他们关系甚少或者毫无关系。[①] 而且，我猜想对经济学家们来说，这或许也是完全无所谓的。（如果世界上的所有金钱都没有物理上的实在性，经济学家有什么理由为此事而操心呢？不管根据哪

① 但对于这有没有关系，我总是拿不定主意。

种理论，绝大多数的金钱都不具备物理上的实在性！）

我提请读者设想一下，坚持我们都是钵中之脑这种古怪的（我还认为它是自相矛盾的）理论的，不是一个与世隔绝的精神病患者，而实际上是某个大国（比如澳大利亚）的全体居民。我们来设想一下，在澳大利亚，仅有少数人相信我们所相信的，而绝大多数人都相信我们是钵中之脑。也许，澳大利亚人之所以相信这一点，是因为他们都是一个头领（悉尼头领）的门徒。也许，当我们与他们交谈时，他们会说："噢，你要是能和悉尼头领谈谈，看看他的眼睛，看出他是一个多么优秀、多么善良和聪明的人，你也许就会相信了。"如果我们问："但是，假若幻觉像你说的那么确切，悉尼头领怎么知道我们都是钵中之脑呢？"他们也许回答说："悉尼头领就是知道。"

如前所述，这不是通常意义上的科学分歧，我们可以想象，澳大利亚人和我们一样善于预见各种经验，善于建造有承载能力的（或看上去是有承载能力的）桥梁等等。他们甚至愿意接受我们最新的科学发现，尽管不是把它们当作真实的东西，而是把它们当作对意象中似乎发生着的事情的正确描述。我们可以想象、也可以不想象，他们在对有关遥远的将来的预测方面与我们有分歧（例如，他们也许会预期某一天那架自动机出了故障，那时人们开始产生一种证明他们的观点是正确的集体幻觉）[①]，但是，他们是否真地作这样的预测，或者他们是否使自己不作任何不同于标准理论所提供的预测，这于我的论证毫无影响。关键是，我已经想象了一种情形，在这里，许多人都有一个与我们极不相同的自成一体的信仰体系。

---

① 如果他们的确作了这样的预测，那么情况就不同了：他们的观点就不再像我们第一章所批评过的那样前后矛盾了，因为他们正在作一个能够（最终）被证明为正当的断言，也就是这样一种断言，它不要求人们去理解一种把真理看作"超验之物"（或无关于正当性辩护）的观点。

这里，没有伦理价值上的分歧问题；澳大利亚人可以如你所喜爱的一样，有一个和我们相同的伦理学。（尽管一位古希腊哲人会说，有智慧就是一种伦理价值，但犹太教和基督教，由于某种救世观念的缘故，事实上都对伦理的概念只作狭义的解释。）

对这些假想的澳大利亚人，我首先要说的是，他们的世界观是狂热的。诚然，“狂热”一词有时几乎是被当作一个褒义词来使用的，此处我并没有这种意思。我认为我们要用极悲伤的心情看待这个坚持如此疯狂的世界观的人类集团。澳大利亚人看来应被称作是心理病态意义上的狂热。把他们的心灵描述为病态的，是一种合于伦理的描述，或者说是近于伦理的描述。但要是不骂他们，人们怎么能够与澳大利亚人争论问题呢？（或者，没法与他们进行争论，因为我将假定他们是不可能被说服的。）

人们马上能想到的一个证据，一定是同这些澳大利亚人的观点的非融贯性有关的，我指的不是第一章中在这种观点里找到的那种非融贯性。那是一种深刻的非融贯性，它需要有一个哲学的（因而是有争议的）论证来加以揭示。而澳大利亚人观点的非融贯性就要肤浅得多了。我们的目标之一，就是应该有办法对我们如何知道我们的陈述是真的给出一个解释。为此，我们一方面建立一种感知的因果理论，以便能从我们的理论本身之内出发，知觉如何来自转换器官对外部世界所施的操作的理论之内，对我们把什么东西当作我们的知觉知识的可靠性作出一个说明。另一方面，为了达到上述目标，我们采用一种关于统计和实验设计的理论，以便能在理论自身之内，说明我们用以排除实验错误的程序是怎样在大多数情况下，的确具有排除实验错误的倾向的。一句话，我们在建立一个关于作为一个整体的世界的理论时，应当包括对我们能据以知道该理论正确无误的活动和过程的说明，这一点是对于我们的理论本身的一个意义重大、极为有用的制约条件。

但是澳大利亚人的那个信念体系并不具备这种融贯性质（至

少在我们看来如此，而且“融贯性”也不是那种我们对之拥有一套算法的东西，而是我们最终凭良心判断的东西）。我们记得，澳大利亚人自己曾设定过一种幻觉，它天衣无缝，以致不存在一种合理的方法，使得悉尼头领能据以知道他所接受并且向其他所有人鼓吹的信仰体系是正确的。依我们的融贯性标准，他们的信仰体系是完全非融贯的。

我们还可以列举一些他们的信仰体系所缺乏的方法论优点。正如我们描述过的，关于在意象中有什么自然规律这方面，他们的信仰体系和我们的是一致的。但是关于那在意象中出现的自然规律是否就是营养钵之外实际成立的自然规律，他们的信仰体系说过点什么没有？如果没有，那就缺乏一种我们所追求的全面性，因为它没有——哪怕用他们自己的话说——指出真正的和最终的自然规律是什么。它无疑违反了奥卡姆剃刀原则。奥卡姆剃刀看来同样很难、甚至不可能形式化为一种算法。但是，事实上钵中之脑理论假定营养钵之外的对象在我们的经验解释中毫无作用，根据这一理论本身，在这里我们可以肯定地说：“若无必要，不要增加实体数目”这一格言已被违背。那些在实质上而不是在字面上遵循奥卡姆剃刀原则的理论，我们称之为功能上简单的理论。

我上面所说的是，我们决定一个科学理论的可接受性的程序必定涉及作为整体的科学理论是否显示某些“优点”。我假定，构造一个科学理论的程序，如果被分解为以句子为单位来证实科学理论的程序，这将是错误的。我假定，科学中的证实是整体性的事情，需要接受经验检验的“法人团体”是整个语句系统，而对这整个语句系统接受经验检验的好坏的判断，或多或少最终是靠直觉完成的。若不对全部人类心理进行形式化，这种直觉的判断便不能加以形式化。但是，我们还是回过来讨论原来的问题：科学中蕴含的价值究竟是什么？

我一直在论证，如果我们采纳我们在批判钵中之脑的过程中

所运用的价值，当然还要加上在这个例子中未有争议的其他价值(例如，我们大概和钵中之脑共同抱有的对工具性效能的欲求)，那么我们就得到了一幅预设了一个丰富的价值体系的科学图景。事实是这样的，如果我们考虑到通过考察科学家和普通百姓把接受什么理论看作是合理的而揭示出来的合理可接受性的理想，那么，我们便可看到，我们在科学中正竭力从事的工作是构造一个具有工具性效能的、融贯的、全面的和功能上简单的等等特性的世界的表象。但为什么呢？

我的回答是，我们之所以需要这类表象，而不需要由澳大利亚人，由钵中之脑们所拥有的"病态"的概念世界，其理由是，拥有这类表象系统是我们人类认知能力兴盛发达的观念的一个组成部分，因而是我们整个人类兴盛、幸福的观念的一个组成部分。

当然，如果形而上学实在论是正确的，并且，谁要是把科学的目标仅仅看作是极力使我们的概念世界和自在世界相"匹配"，那么他也许会满足于说，我们之对融贯性、全面性、功能简单性、工具效能等感兴趣，仅仅因为它们都是达到"匹配"这个目的的手段。但是在我们的表象和自在世界之间的先验的匹配概念是毫无意义的。否认我们需要这种与一个本体世界的形而上学匹配，并不等于否认我们需要与经验世界的普通的经验上的适合(这由我们的合理可接受性标准来判断)。但是，与本体世界相反，经验世界是依赖于我们的合理可接受性标准的(当然，反之亦然)。我们使用合理的可接受性标准来建立一幅"经验世界"的理论图景，然后，由于这幅图景的发展，我们根据这幅图景来修正我们的合理可接受标准本身，如此不断，以致无穷。我们的方法要依赖于我们的世界图景，这一点我在其他书中已经强调过了。我这里所要强调的是这个依赖的另一面，即经验世界对我们的合理可接受性标准的依赖。我现在要说的是：我们即使要有一个经验世界，也必须有合理的可接受性标准，而这些标准揭示了我们的最佳思辨理智这个概念的一部分

内容。简言之，我现在要说的是：实在"世界"是依赖于我们的价值的(反之亦然)。

至少某些价值必定是客观的

和长期流行的观点相反，科学不是"价值中性"的。这一事实当然并不表明"伦理"价值是客观的，或伦理学可以成为一门科学。事实上，无论是把"科学"理解成实验科学还是把它理解成演绎科学，都不存在一种伦理"科学"的前景。正如亚里士多德早已指出过的：[①]

> 当我们谈到此类课题，并且使用此类前提时，我们必须满足于粗略大致地表明真理，当我们谈到这类只是大半为真的事物，并且使用同样类型的前提时，我们必须满足于一些与此差不多的结论。因此，每种陈述都应当按这种态度加以接受，因为对每一类事物的精确性的追求，只求达到该学科本性所认可的地步，乃一明智之士之标志。而接受一个数学家的可能的推理，从修辞学家那里索取科学的证据，则显然是愚不可及的。

但是精确科学(这当然是合理性思维的最典型的例子)中的合理可接受性确实依赖于"融贯性"和"功能简单性"这类认知的优点，这一事实表明，至少某些价值词代表它们所运用其上的事物的性质，而不仅仅表达使用这些词的个人的感情。

如果，"融贯的"和"简单的"这类词并不代表理论的性质，甚至不代表那些模糊的定义得不确切的性质，而仅仅代表一些人对待

① 《尼各马可伦理学》，第1卷，第3章。

理论的"态度",那么,那些用来表达合理可接受性的术语,比如"证明为正当的","得到有力确证的","现成说明中最理想的",也势必是完全主观的了。因为,合理的可接受性不可能比其所依赖的参项更为客观。但是,正如我们在前一章已论证过的,认为合理的可接受性本身仅仅是主观的,这是一个自相矛盾的观点。所以我们不得不得出结论说,至少这些价值词有某些客观的用法,有某些客观的正当性证明条件。

当然,有人会设法否认"融贯的"、"简单的"、"已证明为正当的"等这类词是价值词,从而不承认任何类型的客观价值的存在。有人会坚持说,这些词代表的是我们确实认为有价值的性质,然而我们的这种看法本身并无客观正当性。但这种思路马上会遇到麻烦。"融贯的"和"简单的"与典型的价值词共有的特征数不胜数,如"善良的","美的"和"好的"以及"融贯的"和"简单的"等等,通常是作为赞誉词来使用的。我们对融贯性、简单性、证明为正当的看法,一如我们对善良、美、好的看法,都是受历史条件限制的。关于认识的词和任何美学价值词一样,一直是处在无休无止的哲学争论之中的。约翰·卡迪纳尔·纽曼对合理性的理解明显地与卡尔纳普的大不相同。如果这两人同时活着并且能够会面,其中一方要想使对方信服似乎是很不可能的。到底谁对合理性本身的理解是合理的呢?这个问题之难恰如对一个伦理体系的正当性证明之难。并没有可资求助的中性的对合理性的理解。

有人也许会尝试各种约定论的做法。例如,说"已证明为正当$_{卡}$"和"已证明为正当$_{纽}$"是两种不同的"性质",并且说,在决定用"已证明为正当"这个词去表示"已证明为正当$_{卡}$"的意思还是表示"已证明为正当$_{纽}$"的意思时,是包含一种"主观的价值判断"的,但是在陈述某语句 S 是"已证明为正当$_{卡}$"还是"已证明为正当$_{纽}$"的事实时,并不包含任何价值判断。但是,在使用"事实"一词时,是从谁的立场出发的呢?如果不存在一个人客观上应当具有的合理性

观念，则“事实”概念便化为空无。没有融贯性、简单性和工具效能这些认知价值，我们就没有世界，也没有“事实”，甚至没有某物相对于某物如此这般的事实，因为这类事实与其他事实的处境是相同的。这些认知价值如果不看作是对人类兴盛的整体主义观念的组成部分的话，就是任意武断的。在失去了陈旧的把真理看作“符合”的实在论观点，失去了认为公众“标准”确定证明为正当的实证主义观点之后，我们就只能将我们寻求对合理性的更好理解的努力，看作是受我们的善的观念支配的意向性人类活动，就像其他任何高于习惯和高于仅仅听凭爱好或迷恋的活动一样。

其他领域中的合理性

如果科学，尤其是精确科学中蕴含的价值揭示了我们善的观念的一部分，我想，它的其余部分也能在其他的知识领域里，从我们的合理可接受性标准中发现。当然，在这方面，拓宽合理的可接受标准的概念就是必不可少的事。

迄今为止，我们仅仅从严格的字面意义上考察了合理的可接受性标准，即告诉我们何时应该、何时不应该接受某些陈述。但是广义的合理的可接受性标准不仅仅涉及我们如何判断陈述系统的真或假，而且涉及我们如何判断它们的恰当性和明晰性。因为在某些方面——纯粹认知上的——一个陈述系统除了是假的之外，也可能无法向我们给出一个满意的描述。

如果我曾经作过选择，即使在理论科学方面我也会提出这个观点的。我也许已经指出，精确科学所关注的，不仅仅是发现真陈述，或者发现真的和形式上全称的陈述（定律），而是发现真的和相关的陈述。相关概念带有大量利益和价值。但如果我真这样说过的话，也仅仅是论证我们对世界的知识预设了价值，而并没有提出这样一个激进的主张：被当作实在世界的东西，是依赖于我们的价

值的。

当我们转向知觉的合理性，即转向一个我们赖以决定某人能否对哪怕最简单的知觉事实给出一个真的、恰当的和明晰的解释的隐含的标准和技巧时，我们会看到大量因素起着作用。近来，心理学家强调，即使在最简单的知觉那里也深深蕴含着理论的建构。不仅在神经生理学层次上如此，在文化层次上亦是如此。一个没有文化的人会走进一个房间并对房间作一番描述，但是如果他不知桌椅为何物，他的描述恐怕很难传达其所属文化的其他成员希望得知的关于房间的信息。他的描述也许只包括真陈述，但这种描述是不恰当的。

这个简单例子表明，要求一个描述是恰当的，本身就包含着要求叙述者有一套可行的概念。我们要求合理的叙述者在面对某些描述对象时，能拥有某些概念，并能明了使用这些概念的必要性。如果叙述者未能运用某一概念，我们就有理由批评他以及他的描述。

一个真实的描述，不管其描述对象是房间中的桌椅等简单事物，还是复杂的人际关系和环境，其真实性都不会因此而改变。请看我们在形容他人时每天使用的词，例如，体贴或粗心。这两个词当然会被用来表达称赞和责备。而且，在"事实-价值二分法"的总标题下，还有许多界限已经模糊了的细微的区分，其中之一便是使用语言表达式来描述和使用该表达式称赞或责备之间的区分。但是，单凭词汇本身，是无法作出这个区分来的。某人很粗心，这一判断也许确实用来责备，但也可能仅仅用来描述，也可用来说明或预言。

例如，我可以对你说："别在意琼斯。你可能会认为他那种说话的腔调是讨厌你，其实那只是一种一般的错觉。无论他怎么看你，他都会摆出一副令你难堪的架式。他只是一个很粗心的人，千万不要以为他是专冲你而来的。"

在这段想象的简短话语中，有人使用了"粗心"一词，不是为了

责备琼斯，而是为了向另一个人预言和说明琼斯的行为。这预言和说明也许都是完全正确的。同样，“妒忌”可以是一个责备词，也可以全无责备之意而加以使用。（有时一个人完全有权妒忌。）

“粗心”一词的使用，在我看来是说明实在的世界和实在的语言中极其模糊的事实/价值区分的一个极好例子。艾里斯·默多克在《善的至上性》[1] 一书中已经强调了诸如“粗心”、“冒失”、“固执”等词在实际道德评价中的重要性。即使像“约翰是一个非常粗心的人”、“约翰只想着自己”、“约翰为了钱几乎什么都干”等等陈述，在最实证主义的意义上也都不失为完全真实的描述（注意“约翰为了钱几乎什么都干”这句话不含任何价值词），若要断定这三个陈述的合取，一定就是“约翰不是一个好人”。当我们把事实和价值分开来考虑时，我们一般将“事实”看作是由某些物理主义或官腔十足的术语表达的，而将“价值”看作是由最抽象的价值词（如“善”，“恶”等）来表达的。在事实本身处于“粗心”、“只想着自己”、“为了钱什么都干”这样的层次上时，坚持价值独立于事实就困难得多了。

正如那位应当使用桌椅等概念而未用的描述者遭到我们的批评一样，一个未对某人是体贴的或自发的加以评论的人也同样容易招致批评；说他是麻木、肤浅的；说他的描述是不恰当的。

超边沁主义

现在我们回过来修改我前面举过的“钵中之脑”的例子。这次我们来想象，澳洲大陆所盛行的文化，在历史、地理和精确科学等诸方面和我们相同，只是在伦理学上存有分歧。我不打算提及通常的超纳粹或诸如此类的例子，而是引更为有趣的超边沁主义为例。

---

[1] 《善的至上性》，劳特利奇和基根·保罗出版公司，1920年。

我们来想象一下，澳洲大陆上居住着这样的居民，他们对“快乐主义心境”有着某种精心制定的科学测量手段。并且他们相信，一个人应该永远为求得最大限度的快乐心境（以此来指最大多数人的最大快乐心境）而行动。我假定超边沁主义者特别老谋深算，他们深知预言未来和估计行为后果之困难重重。我还要假定，这些人极其冷酷，并且假定，虽然在有理由怀疑他们的行为事实上能否带来最大多数的最大幸福的情况下，他们不至于借口为最大多数人的最大幸福而使某人痛苦，但是在人们确信其行为的预期后果的情况下，他们会愿意采取最残酷的行为——愿意虐待幼童，以莫须有罪名给人定罪——只要这些行为的后果（在恰当地扣除了无辜受害者的痛苦之后）将会以任何大于零的增量∈（不管多小）增加总的满意程度。

我猜想，我们对这种超边沁主义道德也许不会感到满意。我们中大多数人也许会谴责超边沁主义者有一种病态的价值系统，谴责他们是官僚主义的，或者是残酷的等等。他们是表现形式最为可怕的“新人”。而他们也许会反唇相讥，说我们是软心肠的，是迷信的，是受了旧理性传统束缚的。

我们和超边沁主义者之间的分歧，正是通常为了证明两群人可以在事实方面意见完全一致但在价值方面仍然各持己见而设想的那种意见分歧。但是让我们更仔细地看看这个例子。每个超边沁主义者都熟知以下事实，即在某些时候，最大多数的最大满足（由功利来确定）要求一个人撒谎，而为了极大地提高一般快乐水平撒谎不能视为贬意的“不诚实”。所以下面超边沁主义者对“诚实”这个描述的用法，是完全不同于我们对同一描述词的用法的。同一种说法也适用于“体贴的”、“好公民”等描述词。在对人与人关系的描述上，超边沁主义者和我们各有一套词汇，互相间区别很大。他们不仅缺乏我们的许多描述方式（或者说已把它们改变得面目全非），而且他们很可能发明一些我们所没有的独特行话（例如

描述享乐心境的确切术语)。人类世界的结构将开始改变。随着时间的流逝,超边沁主义者和我们最终将生活在不同的人类世界。

简言之,说我们和超边沁主义者"对事实的看法一致,在价值上存有分歧",这不是事实。在几乎所有的人际关系处境中,我们对事实的描述会完全不同于他们对事实的描述。尽管他们对这些情景所作的描述没有一个是假的,我们还是不会说他们的描述是恰当的、明晰的。同样,他们也不会说我们的描述是恰当的和明晰的。一句话,即使撇开我们"在价值观上的分歧"不论,我们也不会认为他们对人类世界的总体表达是都能合理地接受的。正如钵中之脑无法把握世界的状况恰恰是其病态的合理性标准(病态的理论的合理性标准)的直接后果一样,超边沁主义者无法把握人类世界的状况也恰恰是他们的病态的人类兴盛概念的直接后果。

关于善的主观主义

人们通常认为,从"约翰是体贴的、真诚的、善良的、勇敢的、有责任心的"等等,推出"约翰在道德上是善的",至少包含一个未经证明的(和不可证明的)"前提",即"体贴是道德上善的"。而且据称,要从事实陈述中推出道德结论,先得有一些道德"前提",这说明伦理陈述是不能合理地加以证明的。

把伦理学描绘成一种倒金字塔,它以支撑我们的全部道德信念和道德思维内容的道德公理为尖端(其自身没有支撑),这种说法是幼稚可笑的。还未曾有人成功地把一个公理结构强加于伦理学之上(正如前文引述的亚里士多德的话所指出的,我们能够列举的道德公理差不多总是只能"大部分"为真)。把某人希望怀疑的那些思想,描述为是建立在不可论证的公理之上的,这正是怀疑论者在各个领域都用过的手法。例如,怀疑物质对象存在的怀疑论者论证道:"如果我们感觉到似乎有一个物质世界存在,则很可能有一

个物质世界存在"这个原则是这样一个不可合理地加以支持的前提，凡是我们宣称"观察到"一个物质对象，或者证明对物质对象的信仰的正当性的时候，我们就默默地求助于这个前提。事实上，伦理学、数学和对物质对象的探讨是以概念、而不是以公理作为前提的。概念被用于观察和概括的过程中，而它本身，又由于我们成功地使用它们来进行描述和概括，而获得其存在的正当根据。

那些较老练的反对伦理客观性观念的人，承认我们的伦理信念是基于对特例的观察、直觉和一般准例，而不是基于含糊的道德公理的集合体，但他们又会指责说，伦理观察本身得受一个不可救药的病症——投射的影响。

根据这种解释，人类从本性上讲(尽管是断断续续的)是富有同情心的。因此，当我们看到某些暴行发生，例如，某人仅仅为了虐待狂的快感而殴打儿童，我们(有时)会毛骨悚然。但是"投射"这种心理机制使我们将感情性质体验为行为本身的性质：当我们其实该说"我的反应是毛骨悚然"时，我们通常都说"这些行为令人毛骨悚然"。因此，我们当作伦理观察的内容的东西，实际上只是我们自己主观的伦理情感的观察而已。

这个说法还有更精致的形式(像任何其他说法一样)。休谟曾假定了一种他称之为"同情"的人类倾向，这种倾向在文化的影响下日渐扩展。当代社会生物学家们假定了一个他们称为"利他主义"的本能，并且提到有一种"利他的基因"。但是，关键思想依然如故：存在着伦理情感，但没有客观的价值性质。

我们已看到，这是不正确的：至少存在着某些客观的价值，例如，正当性证明。我们或许还可以宣称，伦理价值是主观的，而认知价值则是客观的；但是，说不可能有任何客观的价值，这种观点已被驳倒。

为表明道德主观主义的观点错在何处，我必须重提在第二章

里用来反对形而上学实在论的论证。你可能会觉得我这样做很不可思议：主观主义不是形而上学实在论的对立面吗？如果谁这么看，那么，任何反对形而上学实在论的论证就好像必定是支持主观主义的；而我准备遵循的策略，即使用这同一个论证既反对形而上学实在论，又反对主观主义，似乎是不可能的事。

但是事实上，形而上学实在论和主观主义不是简单的“对头”。在今天，我们对物理学太倾向于实在论，而对伦理学又太倾向于主观主义。这两个倾向是相互关联的。正因为我们对物理学实在论倾向太重，正因为我们将物理学（或某些假定的未来物理学）看作是“唯一真的理论”，而不仅仅看作是适合某些问题和目的的合理地可接受的描述，我们才倾向于对不能“还原”为物理学的描述采取主观主义态度。对物理学少一些实在论和对伦理学少一些主观主义同样是相互关联的。

第二章结尾处的论证是直接反对形而上学实在论的“物理主义”或自然主义的变种的。为复述它，我们假定标准诠释I（“猫”指猫，“樱桃”指樱桃，等等）与物理主义的关系R或者重合，或者同一。所以R处于符号“猫”（或某人正确地使用这些符号的物理事件）和猫之间。此外，我们所描述的非标准诠释J也与某关系R′有重合，并且可根据R以及用于构造J的可能世界和置换（见附录）加以定义。所以R′处于“猫”这一符号（或某人以标准方式使用这些符号的物理事件）和樱桃之间。R和R′两者都是“符号”：同一些句子在这两个符号之下都是真的。由句子的R′之类真值所引起的行为（即，从主体的观点看，将获成功的行为）与R类之真的句子所引起的行为是相同的。如果R“等同于指标”，如果R、R′和一切以满足我们的操作制约和理论制约的方式将外延分配给我们周围的其他关系不是同等正确的话；如果它们之所以不是同等地正确，是因为其中之一——R——就是指称，那么，从一个物理主义者的角度来看，那个事实本身就是一个莫名其妙的事实。

这个论证不只是反对形而上学的，也是反对还原论的。如果物理主义的世界图景中没有什么是与“猫”指称猫而不指称樱桃这一明显事实相符的，那么，这正是我们所用的一切概念必须还原到物理词汇的决定性理由。因为指称和真理是两个我们不可能前后一致地放弃的概念。如果我认为“一只猫在席上”，这就等于说我应该相信“猫”指称某物（尽管这不是一个形而上学实在论对指称的解释）并且应该相信“一只猫在席上”是真的（尽管这不是一个形而上学实在论对真理的解释）。

回顾了第二章的论证之后，我们来看看它是怎样影响一些支持道德主观论的论据的。“投射”论对道德体验作了这样一种解释：道德体验可以说是错了位的主观感受。将投射论和下列解释作一对比：“人人在一定程度上都有正义感和善心。对于像‘以仁爱之心对待生人。因为你很清楚一个在埃及举目无亲之人处于什么样的境况’这样的要求，我们会（断断续续地）作出响应。我们的同情心之所以增长，部分原因是人们对我们劝说道，同情心是应该增长的。即使我们没有轻而易举或不假思索地发现某受害者是一个我们能够同情的人，我们有时仍会感到暴行是错误的。我们逐渐意识到损人和害己之间、利人和利己之间的相似性。我们为道德上相关的情境特征发明道德用词，然后逐步着手进行明确的道德概括，它导致我们道德概念的进一步完善，如此等等。”

从表面上讲，这个解释较之投射论更为简单，更为精致。（首先，它承认论证在形成道德态度中的作用。）不过，许多有识之士感到，当今时代，我们必须拒绝“正义感”和“有善心”（此处并非主观意义上的用法）之类“不科学”的空谈。于是道德知识变得成问题了，或者干脆是不可能的了。

但是，此处“不科学的”究竟指什么呢？对于有一种叫做正义的东西的信仰并不是一个对鬼神的信仰。正义感也不是一种能使我们体察到这鬼神的超常感觉。正义不是谁想添加到物理学已认识

到的对象的清单上的某种东西，一如18世纪化学家们把燃素列入已被化学所认识的元素的清单上一样。如“不科学”一词所暗示的，伦理学与物理学并不冲突。只是，“公正的”、“善的”和“正义感”属于那类不能还原为物理学话语的话语中的概念。正如我们刚刚看到的，其他一些基本的话语也不能还原为物理学的话语；但并不因此就成为不正当的，对“正义”的谈论，正如对“指称”的谈论，可以是非科学的，而又不是不科学的。

作为一条了解所发生情形的途径，我们应考察任一逻辑和数学的基本原则，比方说，整数数列能无限延续下去（每一位数必有一后继数）的原则，或者，整数的非空集必定包含一个最小的数。假定有人提出下列观点：“这些原则对我们在实践中处理的数和数的集合来说是真的，所以它们似乎是必然的，通过‘投射’机制，我们把这种必然性之感附之于这些原则本身，我们感到这些陈述有一个神秘的‘必然性’。但是实际上，这是毫无正当理由的。据我们所知，这些原则甚至可能是不正确的。”

实际上没有人会同意这一点，相反，实际上每个数学家会这样说：“绝大多数人在某种程序上都有数学的直觉。所以我们直觉地‘看出’或通过例子（或通过巧妙的提问，像柏拉图对话中的奴隶）‘看出’这些原则必然是真的。”

库尔特·哥德尔相信，“数学直觉”类似于知觉：数学对象（他称之为“概念”）是外在的，我们的直觉使我们能理智地感知到这些柏拉图式的实体，但很少有数学家会承诺这样一种柏拉图式的形而上学。哥德尔将数学直觉与知觉相类比，揭示了一个相当简单的知觉观念。视觉并不真正给我们提供抵达现成世界的直接通道，而是提供一个部分由视觉自身构造和建立起来的对象的摹写。如果我们将物理学家所说的虹当作是“自在的”虹，那么自在的虹就没有色带（通过光谱分析产生一个频率的平滑分布）；红、橙、黄、绿、

蓝和紫的色带是知觉之虹的特征，不是物理学家的虹的特征。知觉之虹依赖于我们知觉器官自身的本性，用古德曼的话说：依赖于我们的视觉上的"创世"。（物理学家的"对象"也依赖于我们的创世，对"同样的"对象，物理的极端不同的描述泛滥成灾，便表明了这一点。）但是，我们并不因为视觉在虹里看出了色带而说视觉有缺陷；某个不能看出色带的人的视觉倒可能有缺陷。视觉由其能提供一个和我们相关的客体、而不是一个形而上学的物自体相一致的描述，而被确证为是好的。只要使我们看到现在这个样子的世界，即是部分地由人的视觉本身创造的人和有功能的世界，便是好的视觉。

集合论中一个新提出的公理（比方说"选择公理"）之所以能被接受，部分是因为它符合专业数学家的直觉，部分是因为它的成效。如果选择公理并没有产生堪称卓越的数学成果，那么，某些人发现它是直觉的这一事实恐怕就不会有这样的吸引力了，数学直觉本身是通过数学原则的把握和对证法的遵循来证明和检验的。简言之，数学直觉只有当它能使人们看到"现在这个样子的"——也就是在那个由人类数学实践（包括把数学运用于其他学科）所构造的数学世界之中的那个样子的——数学事实时，才算得上是好的直觉。

一个生理学家或心理学家对视觉的描述不能告诉我们在虹里看见色带是否可称作"正确地"看见。他们对一个人理解数学归纳法原理时的心理活动过程的描述就更不能告诉我们这一原则的正确与否。一旦明了这一点，对于以下情况，即当一个人"看出"一个行为不正当时，对他的脑运动过程的描绘并不能告诉我们这种行为是否确实是不正当的，就不会大惊小怪了。

对道德"知觉"的讨论，就像对数学直觉的讨论或对指称和理解的讨论一样，是不能还原为物理学的语言或世界图景的。这并不意味着物理学是"不完备"的。物理学可以是完备的，即相对于物理

学的目的而言是完备的。物理学所缺乏的完备性乃是全部特殊理论、图景和话语所缺乏的那种完备性。因为没有什么理论或图景对一切目的而言都是完备的。如果伦理学不可还原为物理学就表明价值是投射，那么色彩也是投射。自然数也是如此，就此而论，“物理世界”也是如此。但是，在这种意义上是投射，并不等于就是主观的[①]。

权威论和多元论

至此，我所论证的是，只要有一个世界，就得先有一个合理的可接受性标准。这个世界，要么是一个“经验事实”的世界，要么是一个“价值事实”的世界（一个有美和悲剧的世界）。很显然，我们不可能既有一个合理的可接受性标准而又不去接受它，或者说与它保持一段距离。（拒绝有任何合理性标准的那种怀疑论使得一个人根本就没有任何概念，正如塞克斯都·恩皮里柯所提出的，这种经验论最终不能用语言表达。）正如我们有充分理由把某些“认知的”气质看作是病态的一样，我们也有充分的理由认为某些“评价的”气质是病态的。但这样说并不等于拒绝多元论，或者说使自己投向权威论。

即使在科学中，主张科学是一项客观的事业（根据公认为人类中心论的“客观性”标准，然而正如韦金斯曾经说过的，这是我们可

① C·S·刘易斯在《人的废除》（麦克米伦出版公司，1947年）中曾提到过投射论的一个无意但又有趣的变种。刘易斯援引一本中学英文教科书（出于宽容，他未指名）：“你记得，有两个旅游者出场了（刘易斯正在讲著名的科尔里奇在瀑布边的故事）：一个说这瀑布‘壮观’，另一个说它‘优美’。科尔里奇心里赞成第一种判断，不喜欢后一种说法。盖厄斯和蒂特斯（刘易斯为未点名的教科书作者杜撰的名字）作了如下评论：‘当这个人说瀑布是壮观的，他似乎是对瀑布作一个评论……事实上，他不是在评论瀑布，而是在评论自己的感情。他真正说的是：我心里有一种与壮观一词相联系的情绪。或更简单地说，我有壮观感。’”

能有的唯一标准），也不是主张每一科学问题都有一个确定的答案。有些科学问题也许有一些客观上并非确定的答案，也就是说，哪怕在科学探究的理想极限上也不存在朝着这些问题的一个答案的会聚；而有些科学问题也许有一些确定的、但与语境有关的答案（例如，约翰心脏病发作的原因是什么？这个问题由不同的人抱不同的目的提出，就会有不同的正确答案）。同样，主张伦理探究在以下意义上是客观的，即某些"价值判断"肯定是真的，某些价值判断肯定是假的，更一般地说，某些价值观（和某些"意识形态"）肯定是错的，而某些价值观则肯定是低于其他价值观的，这与愚蠢地坚持根本没有任何不确定的情形完全是两回事。（伯纳德·威廉斯着重指出了一种特别重要的不确定的情形，在这种情形中，所有的选择都很可怕，没有一个完全合理、明智的人会取其中任何一项。）而在伦理学中，有着语境相对性，这一点自不待言。

如果说今日之我辈已不同于亚里士多德，那实指我们与亚里士多德相比，更具有多元的态度。亚里士多德认识到，不同的幸福观念，不同的人类兴盛观念，由于人们素质的不同，或许是适合于不同的个人的。但他似乎认为，理想地说，有某种人人应有的素质。他认为，在一个理想世界中（那类由谁种田谁做面包之类的琐事暂且忽略不计），人人可能都是哲学家。我们同意亚里士多德的观点：不同的人类兴盛观念适合不同素质的个人，但是我们进一步相信，即使在那个理想世界中，也可能存在不同的素质，我们还相信，多样性也是理想的组成部分。我们看到，在不同理想之间有某种程度的悲剧冲突，某些理想的实现永远要以牺牲另一些理想的实现为代价。但是，我再强调一遍，相信一个多元化的理想，不等于说相信每一个人类兴盛的理想都与其他理想同样美好，我们拒绝错误的、幼稚的、病态的、片面的人类兴盛的理想。

承认伦理客观性也不应和承认伦理的或道德的权威论相混淆。这是截然不同的两码事。或许正是因为这种混淆，使得一位杰

出的哲学家[1] 赞成他自己所说的一种褊狭的"非认知主义"，并且说"对每个人来说，'最充实的人生'是什么这个问题的最终权威必须是那个人自己"。（注意"最终权威"一词的含混性：这究竟是指政治上的权威，还是指认识论上的权威，抑或是像他使用的"非认知主义"一词所暗示的，指不存在既成事实？）对作为自主的道德主体的人的尊重，要求我们尊重他人为自己选择道德标准的权利，而不管我们觉得他们的选择是多么令人反感。按照政治自由主义的哲学，它还要求我们坚持政府不应通过制定国家宗教和国家道德来预先左右个人的道德选择。但是誓死反对形形色色的政治上和道德上的权威主义，并不应该使人投向道德相对主义或道德怀疑主义。政府为公民个人指定道德之所以是错误的，并不是因为在有关什么生活方式令人满意、什么生活方式不令人满意，或某些其他方面在道德上是错误的问题上，不存在既成事实。（如果没有道德错误这类东西，则政府强加道德选择也就不是错误的了。）许多人为之担心的一个事实是，如果他们公开承认任何类型的道德客观论，他们就会发现某个政府正在用它自己的道德客观性概念来堵他们的口。毫无疑问，这正是为什么有如此众多的人赞同一个他们未必真正同意的道德主观论的原因之一。

---

[1] 见大卫·韦金斯："真理、发明和生活的意义"，《不列颠学院会议录》，第62卷，1976年。

第七章 理性和历史

随着科学的兴起，人们开始意识到，许多问题不能用精密科学的方法来解决，意识形态问题和道德问题是最明显的例子。而且，随着我们对物理学家、宇宙学家、分子生物学家的崇敬和尊重的增长，对政治思想家、道德学家、经济学家、音乐家和精神分析学家等等的尊敬和信任则随之降低。

在此形势下，有些紧跟文化潮流的人主张，在精密科学（和成功地效法了精密科学的社会科学，仅此而已）之外，无所谓知识。这种观点可能采取实证主义或唯物主义的形式，或者采取这两者的结合形式。另有一些人极力证明，科学也是"主观的"和任意的。人们在读库恩那本极为成功的《科学革命的结构》一书时，通常就是这么理解的，尽管库恩现在说这不是他的本意。还有一些人，例如马克思主义哲学家和宗教哲学家，采纳了某种复式簿记法，把技术问题留给精密科学和工程学，把意识形态或伦理问题留给另一种法庭：留给政党，乌托邦似的未来和教堂。但是对以实证主义形式或唯物主义形式出现的极端科学主义立场，对主观主义和极端相对主义立场，对无论何种复式簿记法的做法，几乎没有人会感到满意。正因为我们感到不能满意，这个领域里才存在着一个真实的问题。

确实，这个问题从某种意义上讲是不真实的。一个主张伦理和政治观念是不可证实的人同时又竭力为他自己的伦理观和政治观点作论证。休谟说，只要他离开书房，就摆脱了怀疑论。相对主义

者很可能也会说只要离开书房，就摆脱了相对主义。但这仅仅表明，没有人能够一以贯之地靠相对主义过活。如果这就是回答相对主义时所能说的一切，那么我们只是从相对主义推到1945年那种样式的存在主义（“这是完全荒谬的，但你不得不选择”）。当真是这样不同吗？

为使我们的观点更明确一些，我们来回忆一下上个世纪的一位哲学家，他的功利主义实际上包含了大量的相对主义。我想说的是边沁，是边沁的一个富有挑战性的论断：“撇开偏见不论，针戏游戏和艺术、音乐科学以及诗有同等价值。”撇开偏见不论，针戏和诗是同样好的。

对现代读者来说，这种观点之所以耸人听闻，是因为它和我们通行的文化价值有尖锐冲突。在我们的时代，艺术已经被我们高扬，其地位比它在柏拉图时代和中世纪要高得多。正如许多论者已指出过的，在今天，对某些有教养的人来说，艺术就是宗教。即是说，艺术乃获救的最切近可行之途。

边沁说的是，偏爱音乐和诗的艺术和科学甚于针戏这类儿童游戏纯属主观，正如偏爱香草冰激凌甚于巧克力冰激凌。他并不想否认音乐和诗比针戏游戏确实更有价值（“撇开偏见不论”是这句话的一个重要部分）。放到他的功利主义的理论背景中，大多数人确实宁爱音乐诗歌而不爱针戏，这一事实赋予音乐和诗以更大的“功利”，因而，也就赋予它们以更大的价值。然而价值可以说是“偏见”的产物（即是说纯粹是主观兴趣）。撇开人们宁爱诗而不爱针戏这一事实不论，在针戏和诗的相对价值方面没有什么既成事实可言。边沁说，我们不是因为诗比针戏有更大的价值，才偏爱诗而不偏爱针戏的，而是相反，因为人们偏爱诗而不是针戏，诗才比针戏更有价值。（显然没有任何理由。）

这样粗略地道出上述立场已经使它变得难以置信。现在让我们来考察一个真正“主观的”偏好。

人们似乎时常挂在心上的关于主观偏好的典型例子是这样的。有一种东西C是巧克力冰激凌的滋味，有一种东西V是香草冰激凌的滋味，有两种感情L、D分别代表喜欢和讨厌。当琼斯喜欢香草而史密斯讨厌香草（史密斯喜欢巧克力，琼斯讨厌巧克力）时，所发生的全部事实便是，只要琼斯吃香草，便体验到V+L，并且只要他吃巧克力便体验到C+D；而当史密斯吃香草时，他便体验到V+D，并且当他吃巧克力时便体验到C+L。

但是正如柯勒早已指出过的，这种解释从心理学上讲是浅薄的。香草的滋味如何，对喜欢香草的琼斯和不能容忍香草的史密斯来说是不同的。因此最好是这样：称香草所具有的、只适用于史密斯的性质为Vs。Vs是一种"不愉悦"的味道。可以想象一个人可能体验到Vs并且喜欢它，但这种事极少发生。即使这个人果真喜欢它，也可能会有某种感情上的分裂或压抑。简言之，从心理学上讲（如果不从形而上学上讲的话），Vs是"内在地"不愉悦的，当史密斯吃香草时，他感觉到D（讨厌香草的味道）。因为对他来说，香草有味觉性质Vs。同样，只适用于琼斯的香草的味觉性质Vj是内在地"愉悦的"（这是琼斯感到L，即喜爱的原因）。套用摩尔的话说，滋味Vj和肯定的价值在琼斯那里是有机结合在一起的，滋味Vs和否定的价值在史密斯那里是有机结合在一起的。从现象学上讲，它们不能真正地像"Vs+D"、"Vj+L"这种记号所显示的那样，分成两个部分。差不多可以肯定地说（除了压抑和分裂这些特殊因素之外），如果香草冰激凌在史密斯口中引起Vj而不是Vs，他也许会喜欢香草冰激凌，同样，如果香草冰激凌在琼斯口中引起Vs，他就会讨厌香草冰激凌。

那么，何以我们称对香草而不是巧克力的偏好是"主观主义的"？我的意思是说，即使我们并没有认为所有的价值判断都是主观的，或同意边沁的观点（撇开偏见不论，针戏与诗是同样好的），我们为何还把这种偏好认作是主观的？很明显，如果我们把所有偏

好都认作是主观的，我们将这个偏好也认作是主观的。但是饶有兴味的问题是，为什么这个判断对我们来说看上去不像是客观的(除非我们*就是*史密斯或琼斯)。为什么它不具备许多价值判断肯定具备的那种客观性呢?

这里不仅是有一个分歧的问题。如果我们认为毕竟还有客观的(或者有根据的)价值判断的话，我们很可能认为某些热烈争论中的判断是客观上正确的。纳粹对"*仅仅由于种族原因而大肆杀害犹太人是错误的*"这一判断持有异议，但是反纳粹者并不认为他们与纳粹在这个问题上的分歧是"主观的"。那种认为同性恋在我们的社会中应该完全合法的观点与认为同性恋行为和同性恋者的人权应该从法律上加以禁止的观点是截然相反的，但争论的双方都否认自己的立场是"主观的"。的确，分歧通常使人们更加确信自己的道德立场是正当合理的。因此，使得史密斯/琼斯在偏好中的分歧成为主观的，不只是有人喜欢巧克力、有人喜欢香草这一事实。

事情的部分真相也许是，大多数人在香草和巧克力之间没有强烈的偏好。但这还不是至关重要的。如果有半数人不能忍受巧克力而喜欢香草，另一半人则不能忍受香草而喜欢巧克力，我们也许依然(如果我们对自己的偏好还有理智的话)可以将此称为"趣味问题"。也就是说，把它当作主观的。起关键作用的并不是中立者的存在。

在我看来，确实至关重要的是，造成史密斯和琼斯的爱好的生物学和心理学上的无论什么特征，都是与心灵和性格的重要特性毫无关联的。如果我们设法做这样一个假想实验，想象有一种性格，无论是就其自身，还是就它对情感、判断或行为的影响而言我们都称它为坏的性格，此外又想象人人都知道好的心灵和性格永远在对香草的偏好中表现出来，而坏的心灵和性格则在对巧克力的偏好中表现出来，那么，我认为我们会发现，我们越是使得这种情况对我们显得逼真，我们越会感到，在这样的世界里，前一种偏

好会被视作是“正常的”、“正确的”，而后一种偏好是“反常的”、“邪恶的”，或诸如此类的东西。

我并不是要主张，对所有的偏好都通过据认为是由它们所展现的性格而从道德上加以评判。某些“偏好”本身异乎寻常地重要：谁把为寻开心而虐待幼童视为一桩乐事，谁就有可能（如果他并不是说着玩的话）为这种态度而受到谴责。但是如果所偏爱的事本身并不重要，那么我们是就偏好产生争论还是将它当作是“趣味问题”，通常将依赖于我们认为这些偏好说明什么问题（如果这些偏好果真说明什么问题的话）。价值判断通常是一组一组地出现的。价值判断通常表达了心灵、人格性和性格的持久特性。“我宁要香草而不要巧克力冰激凌”是独立于任何这类有兴味、有意义的价值判断组的。使得个人口味成为主观的东西的，正是这种独立性（当然，还要加上这选择本身缺少任何内在的重要性这一条）。

即使史密斯对香草的偏好是“主观的”，这偏好也不会因此而改成不合理或任意性的。史密斯之喜欢香草有一个理由——可能的理论中最好的那个，就是他感觉到的香的滋味。如果把“主观的”理解成“相对的”的话，价值可以是主观的，同时又不失其客观性。对史密斯来说，香草吃口要比巧克力好，这一点是客观的。在《善的至上性》一书中，默多克指出，法国存在主义者和逻辑实证主义者这两类区别极大的哲学家实际上拥有一个共同的价值判断模式，即认为理性为心灵提供中性“事实”，而意志必须在这些事实的基础上任意地选择“价值”——；价值选择必须是任意的，恰恰是因为事实（由定义可知）是中性的。但是，既然理性对意志如何进行判断没有提供任何线索（根据这种说法，理性只提供事实），那么意志在任意选择时就没有任何理由，正因为这样，法国哲学家称选择为荒谬的，那些自然主义倾向较强的哲学家将直觉和情感（这些是边沁的适用于各种用途的愉悦范畴的历史后继）视为道德选择的终极基础。

在我适才考察的例子中，存在主义-实证主义模式并不适合。“事实(味道本身)”和“价值(味道的好等等)”至少在心理学上讲是同一个东西。一般说来，已呈现的经验性质不是中性的，它们好像经常需要对其有所反应或持有某种态度。有人也许会把被感受到的对正当和充足的理由的要求弃之不顾，正如一个孩子为了从免疫剂中获得益处而学会忍受注射之苦一样，但是特殊经验的明显的好和坏是不容否认的。(相当有趣的是，对这一点柏拉图和中世纪人是承认的——我们或许是最早把经验视为中性的人。)

经验的非中立性也出现在针戏/诗的例子中。我们发现，很难想象一个真正的诗歌爱好者，一个能分辨诗之上品和劣作的人，一个能点评宏伟诗篇的人，居然会喜欢孩童的游戏甚于喜欢给生活增添色彩的诗和音乐等艺术。我们有理由更喜欢诗而不是针戏，其理由在于宏伟诗篇的感受经验，在于宏伟诗篇以后的效果。这就是，通过对我们的想象和隐喻的各种能力的增强，通过诗的想象和隐喻与我们和诗人长期生活在一道的世俗的感受和态度两种不同的经验的结合，我们的想象力和感受性得以强化。这些经验也是明显好的，或者用一个旧式的用语说，它们不仅仅是好的，而且是高贵的。

当然价值判断有其理由——真正是为特殊个人作特殊的价值判断的好的理由——并不意味着，所有价值判断都是合理的。价值判断，人们热切关注的这种判断，人们曾以其名义杀害和虐待他人的这种判断，通常建立在一种攻击性冲动和自恋性观念的不健康混合体的基础上。毫不奇怪，当一个像福科这样的相对主义的历史学家或哲学家撰写历史时，他常常将我们的注意力集中在这些不合理的观念和价值判断上。但是重要的是要明了他这样做的理由何在。

福科研究的是近代初期(16 和 17 这两个世纪)和一般意识形态及文化。他关于事实的知识富于传奇色彩，尽管有许多专家怀疑

福科的“事实”。他的某些著作高度抽象(例如《知识考古学》),而有些则相当具体,如《诊所的诞生》、《监狱的诞生》。《诊所的诞生》也许是福科最好的著作,并且成为福科更抽象理论的例证中的一个重要部分。

福科企图表明,“诊所”,即医院和相关的医疗机构,反映了关于疾病和健康的某些意识形态取得了能和科学知识和技术的进展相媲美的增长,这种意识形态接着又和更为广泛的意识形态的变化相联系,特别是和17世纪个人主义的增长相关。他认为,诊所并不是治疗大多数病人的最好途径,我们对医院作用的信念只是一种意识形态的偏见,一句话,是一种愚蠢的观念。

由此引出的更进一步的看法是,意识形态方面的信念和相关的价值判断是一件相当任意的事情。①在意识形态方面毫无客观性可言(当然,福科自诩的客观的“知识考古学”的神秘出发点除外)。

要知道福科的真实意图,我们来考察一个更熟悉又较少争议的例子。在中世纪,人们相信君主政体是政府的自然的、正当的形式。这种信仰部分建立在现在看来是没有根据的事实信仰上(即认为民主政体会不可避免地产生无政府主义和专制暴政),部分建立在教会的权威之上。教会的观点事实上部分地出于政治上的考虑(教会是国家宗教),但是由于教会本身被认作是上帝的意志和旨意的神授圣钦的解释者,这一点不为人所察。福科想说的是,我们在最近的过去以及(根据其言外之意)在当下的现在所拥有的信仰,并不比对中世纪神圣王权的信仰更为合理。

---

① 由此引出的另一个看法是,我们的思维正是被我们使用的语言所决定。福科指出了“不为我们所察地决定着我们最随意举动的潜在系统”(见J·西蒙:“与米歇尔·福科的谈话”,载《党派评论》,1971年,第2期,第201页)。法国结构主义(至少是福科、阿尔图塞、拉康和德列兹等人阐发的结构主义)似乎通常等于(1)决定论;(2)相对主义;和(3)宣称结构主义是“语言科学”。

我们暂且来考虑一下对神圣王权的信仰，如果我们认为没有正当的理由去相信人格上帝的存在（这个上帝指挥着我们以某种社会形式和结构生活着），那么，这个信仰就会立即被冠之为“不合理的”（这并不否认该信仰可满足真正的心理学的需要）。即使我们相信上帝，如果我们不相信教会有领悟上帝旨意之特殊门道，我们也会认为神圣王权过去和现在都是不合理的原则。说到底，甚至天主教的信徒也会承认，中世纪教会对君主政体的支持的基础是政治上的考虑，其重要性不下于大君和正统神学。简言之，对神圣王权的信仰缺乏（而且从来就缺乏）一个适当的合理的基础。

那么，这种信仰是如何产生的呢？通常的回答或许是部分地求助于政治因素和经济因素（不一定只有成为马克思主义者才能承认这些因素属于意识形态的决定因素），部分地求助于心理学因素。这种由信仰一位人格上帝和来世而来的安慰是明显的，相信一贯正确的教会和神定的社会秩序中的人所得到的安慰大概也是显而易见的。简言之，自恋的自我满足和社会调节是产生这种信仰的真正决定因素。而且，如果这种信仰是真正有代表性的，如果它真是我们全部“意识形态”信仰的代表，那么这些因素就是全部意识形态真正的决定因素。

正因为相信上述结论，许多现代法国思想家对马克思、弗洛伊德和尼采评价甚高。马克思、弗洛伊德和尼采在有一点上是相同的，这就是，他们把我们珍爱的宗教和伦理观念看作是不合理之物的反映；看作是阶级利益的（在马克思那里）、无意识的（弗洛伊德和尼采）和权力意志的（尼采）反映。在我们欣然引为深奥无比的精神和道德洞见的东西下面，潜藏着一口权力角逐、经济利益和自私幻想的沸腾的大锅。这个观点是当今相对主义的利刃。

当此同时，没有一个相对主义者能以我刚才描述相对主义观点时的用法来使用“不合理的”一词。这种用法是被相对主义自身排斥的。我曾把这些段落给一位持相对主义观点的朋友看，他对我

关于神圣王权的文字表示不满。是我没有意识到有识之士已经过哲学家论证而相信这个教义是正确的？我是在兜售一个廉价的马克思主义的或弗洛伊德主义的观点吗?在过去,对神圣王权的信仰当然是合理的。

我对他的回答是:“合理的”可能有这样一种意义,根据这种意义,任何能理智地、有说服力地从一个文化的共同的假定中获得辩护的观点都是一个“合理的”观点。但是那个意义不可能是唯一重要的或按常规理解的重要的意义。犹太人立摩西为立法者和先知,因为他的教义满足真正的宗教需要、文化需要和民族需要。这与通过合理的论证而相信是两码事。后来,在极力劝说犹太人不设国王无效之后,先知给犹太诸王行了涂油仪式。这证明不了后来的基督教国王是神定的。继承了犹太教《圣经》的基督教成了罗马帝国的宗教,这不是因为罗马帝国的臣民和帝王有合理的证据说基督教是正确的。然后,和犹太国王一样,罗马皇帝也由教会涂了油。这证明不了罗马皇帝是神定的,最后,当基督教的假定被接受以后,人们就根据这些假定对神圣王权作“合理推论”。但是,要出于这种考虑而说在罗马帝国后期或中世纪,“对神圣王权的估价是完全合理的”,就贬低了合理性概念。

引进了“理性”自身在历史中变化这一思想的黑格尔,曾讨论过两个合理性的概念。其含义之一是,所谓合理的是由一既定时间内“精神”在历史发展过程中所达到的水平来衡量的。在诸如此类的意义上,有人宣称,对“神圣王权的信仰在当时是合理的”。在黑格尔体系中还有一种合理性的极限概念,它指的是一种注定为恒定不变的东西,是其自身不能被超越的“精神”的最终自我意识。当现代相对主义者通过抛弃这个真正合理性的极性概念而将黑格尔“自然化”时,他们把这一学说变成了一种自拆台脚的文化相对主义。

但是,没有哪位相对主义者想对一切都持相对主义态度。这些

法国思想家们是如何使他们自己的相对主义有所限制的呢？答案因人而异。像阿尔图塞这样的马克思主义者，所持的观点是"阶级利益论"的翻版。一切"意识形态"都是非理性因素的产物，但只有那些是工人阶级利益的产物的意识形态（在现阶段）是"正义的"，并且代表人类解放的方向，而那些出自剥削阶级利益的意识形态是非正义的，并且会带来灾难。但是阿尔图塞将自己的观点和早期的阶级利益论者的观点作了区分。因为他甚至否认，马克思主义的意识形态（工人阶级的意识形态）是真的，或比资产阶级意识形态更接近真理。在阿尔图塞看来，意识形态只能是正义或非正义的，不能是真的或假的。[①]（他说"真"和"假"适用于"实验科学"，并且可能适用于那些有明显经验检验条件的普通的经验陈述。）福科在其近著中似乎也在向阶级利益论靠拢，尽管眼下还未有定论。这种观点（至少就其极端的阿尔图塞形式而言）的要旨在于，它企图维护一切意识形态都不可能是合理的这种极端相对主义的观点。同时，又通过以不同于合理的可接受性的根据来区分好与坏、正义与非正义的意识形态，以保全有些意识形态（在阿尔图塞那里，首先是指马克思列宁主义）可以是好的这种观点。所谓有些意识形态是好的是指，尽管所有的意识形态都是由不合理的或非理性的原因而接受的，但有些非理性的原因（工人阶级利益）是好的，并产生好的意识形态（根据定义？）；而有些非理性的原因是坏的，并产生坏的意识形态。我们不应该依其理由（它们永远是合理化的根据），而应依其原因来判断意识形态。

---

① 根据阿尔图塞的观点："哲学命题是论题。""哲学论题，就其不能在严格科学的意义上加以证明（即像数学或逻辑学中的演算证明那样），也不能在严格科学的意义上加以验证（即像实验科学中的验证那样）而言，可以否定性地看作独断的命题……既然哲学论题不能证明也不能科学地验证，就也不能说是'真'的（像在数学和物理学中那样被证明或被验证）。哲学论题只能说是'正当的'"，载《哲学和学者自发的哲学》第13—14页，马斯佩洛出版社，1967年。

然而，用于限制自己的相对主义的这种途径显然是行不通的。工人阶级利益的胜利会导致一些显然令人向往的结果，会导致一个没有战争和种族主义的世界，而不是导致一个极权主义和打着“社会主义”旗号的帝国主义，作出这样一个判断的根据何在？如果答案是，后者不是“真正的”社会主义，或者它没有反映工人阶级的真正利益，那么，说某一特殊组织（如法国共产党，阿尔图塞是其主要成员之一）或政策将会促进工人阶级的真正利益并且代表“真正的”社会主义这一判断的根据又何在？如果这些信仰能被合理地加以证明，则并非任何意识形态都是不合理的；如果不能，那么说某一组织或政策是“正义的”这一论断，就和所有其他的意识形态一样，必定是完全不合理的。如果人类所有关于意识形态问题的思想都是损人利己的愚蠢念头，那么关于哪种信仰是来自工人阶级的利益、哪种不是的思想，也必是损人利己的愚蠢念头。

还是回到福科身上来。撇开他的近著中所表现出的激进化的迹象不谈，他专注于他选择的事件的动机**恰恰**是想指出人们在接受意识形态立场时所特有的真实**理由**是完全非理性的（甚至是不合理的）。这里的意识形态概念相当宽泛，它不仅仅指共产主义、法西斯主义、民主、神圣王权等这些正在讨论的概念。某人生了病，需要照料，某人是个罪犯（如有可能）应加以“归化矫正”，这些信念以及我们许许多多的日常信念，在这些思想家看来都属于“意识形态”的范畴。的确，在马克思主义社会学家和法国哲学家的锐利眼光中，几乎每一信念都属于“意识形态”的范畴。或许，“如果我扔掉眼镜，它会摔碎”这句话在意识形态上是中立的，但是，这种情况犹如凤毛麟角。

或许我们遗漏了福科所言之精义。他自己恐怕会说，他的实际意思并非认为既往的意识形态观念是愚蠢的或不合理的，而是在于，在他赋予这个词的极为宽泛的意义上的一切意识形态，包括我们当前的意识形态，都是文化上相对的，他力图表明，人类每一种

文化中的人是如何在一系列包括若干非理性决定因素的、无意识的先入之见的支配下生活、思考、观察和做爱的。如果先前的意识形态如今已显得“不合理”了，这是因为我们使用受制于我们文化的合理性概念对它们进行判断的缘故。

福科这一解释的令人困惑之处在于，他和其他法国思想家所列举的决定因素，用我们现在的眼光来看是不合理的。如果我们现在的意识形态是以自己的眼光认定的不合理的力量的产物，那么这种解释就是内在地前后矛盾的。法国思想家不仅仅是文化相对主义，他们还从内部攻击我们当前的合理性概念。这正是读者所感受到、并为之困惑的地方。

文化相对主义根本称不上是一种新的学说，人类学家自有人类学起就在向我们鼓吹文化相对主义。但是将福科的相对主义等同于旧的相对主义恐怕是一个错误。

人类学家在向我们鼓吹相对主义时通常援引一个异族部落的信仰和惯例。这些信仰和惯例一上来会使我们产生一种强烈的不合理的或令人厌恶的、或两者皆有的印象。随后，他要说明，这些信仰和惯例实际上促进了福利和社会内聚力。简言之，在他援引的例子是合理的情况下，他要向我们表明，那些在我们的社会中被认为是错误和不合理的信仰和惯例，在不同的自然和社会环境中也许会被认作是正确和合理的。

当然，人类学家通常从他们自己的例子中(有些例子远没有人类学家想象的那样清晰)得出错误的结论。人类学家经常会说：“这完全是相对的，”意思是说，不存在任何既定事实以决定孰是孰非。理查德·博伊德曾对我道出个中原委：他们这样做通常是出于政治上的动机；他们是要通过攻击我们对我们自身的优越的合理性和道德的信仰，从而说服我们不再摧毁原始文化。不幸的是，这种论证是将不同的问题混为一谈。人类学家举的例子(如果例子还不错的话)只说明，对和错是相对于环境的，而不是说，根本无所谓对

错，即使在特定的环境中也是如此。他反对文化帝国主义的推论只等于说：外域文化客观上并不比我们的文化更差（因为在他看来，没有客观的好和差这类东西）；因此这些外域文化和我们的文化恰好是同样出色。

这个推论值得怀疑。结论要求"恰好和……同样出色"指客观上恰好和……同样出色（至少以我们的眼光来看）。但是从不存在客观的价值所推论出来的，不可能是某一事物（在所要求的意义上）"恰好和其他事物一样出色"，而是没有"恰好和……同样出色"这种东西。如果价值真的是任意的，那么为什么我们不去摧毁我们所喜爱的任何文化呢？

幸运的是，较之否认客观的价值，批评文化帝国主义的理由更为充足。人类学家的动机也许是好的，但他选择了一个错误的论据。另一个他会含糊其词的概念是"相对的"。他的例子实际上肯定的是杜威的"客观相对主义"。某些事情在某种境遇下是正确的（客观地正确），而在另一种境遇是错误的（客观地错误），而相关的境遇则是由文化和环境构成的。关于这一点，人类学家是正确的，但这与说价值的相对性只是指价值是观点和趣味的问题完全是两码事。

另外，摆脱了概念上的混乱，我们就不会为人类学家的推论所困惑。我们会欢迎他们的考察，因为它会拓宽我们的感受性，并且打击我们的文化至上性的自命不凡的假定。但是将福科的推论和人类学家的推论作一比较，即可看出他们的差别。福科所论证的，不是过去的惯例比它们现在看上去的更为合理，而是在论证所有的惯例都不那么合理，实际上都主要是由不合理的、自私的力量所决定的。这个学说与从前的文化相对主义只有着表面的相似性。

事实是，我们一直在讨论的观点迎合了这样一种理智诱惑，这种诱惑是我们对于心理学机制和社会学机制的增长了的知识和感受性的产物。这种知识和感受性真假参半。这种诱惑使人陷入这

样的结论，认为所有的合理论证仅仅是合理化，然后又使人着手为这一立场作合理的论证。

如果所有的"合理论证"都仅仅是合理化，那么，不仅试图为任何观点作合理的论证是没有意义的，而且坚持任何观点也是无意义的。如果我把自己的赞同和反对看作是疯狂的举动，那么我便会中止赞同和反对的举动。——这种不可能合理地赞同和反对、而只能无理智地、拙劣地摹仿合理讨论的东西是不能称之为"陈述"的。恩皮里柯最后得出结论说，他自己的怀疑论也不能由一个陈述来表达(因为他甚至不能知道像"我不知道"这一陈述)。和恩皮里柯一样，现代相对主义者，只要他还是前后一致的(但是一个人如何可能前后一致地坚持一个使得一致性概念变成无意义的学说呢?)，就应该以他自己的言论只是情感的表达这一结论告终。

这样说并没有否认我们能合理地、正确地认为我们的某些信念是不合理的。这是说，坚持我们在理智方面都无可获救的主张要能成立而不至于不可理解，是有一些限度的。例如，我们的确讨论了像福科所提出的那些学说；我们力求公正无私；我们极力接受波普尔所说的"批判态度"，并且积极寻找我们可能忽视的证据和推论，哪怕这些证据和推论会和我们自己的观点作对。如果我们不认为这些讨论和交际的惯例以及这些批判和公正无私的品质会瓦解(如果不是立即地，那么也会逐渐地，直至最终地瓦解)不合理的信仰，并会加强我们的最终结论的有根据的可断定性，所有这一切都不会有丝毫的意义。合理性可能不是由一个"推测"或一套原则来定义，但是，我们的确拥有一个关于各种品质的发展着的观念，用以指导我们的行动。

有人会反对说，这个观念只是老调重弹，没有什么新东西。卡尔纳普和纽曼都是谨慎而有责任心的思想家，他们都承认刚才提到的认知品质，但是，假若他们生活在相同时期并有幸相遇的话，他们都会认为不可能说服对方。但是没有一条人人满意的解决一

切争论的途径，这一事实并不表明，在这种情况下便没有好差之分。我们大多数人认为，纽曼的天主教多少是有些令人困惑的；而大多数哲学家认为，尽管卡尔纳普才华横溢，他的某些论证也经不起推敲。我们作这些判断，这个事实说明，我们确实具有一种关于公正的、倾听对方意见并且不偏不倚的理智的调节性观念。我们确实认为，在这个或那个方面思想家为什么及如何达不到该理想是有确定事实的。某些人会说："那又怎样呢？如果没有一个在我们正在争论的观点之外的、我们可以诉诸的合理的可接受性概念，那么，就算我们最终解决了一个实际的争端，我们的处境也不会由此改观。把它运用于一个不可调和的争端（诸如刚刚假想的卡尔纳普-纽曼的争论），这种说法固然不错；但是，如果我们放弃了确有公正性、一致性和有理性这类东西存在这一观念，如果我们对能在自己的生活和实践中接近它们也不承认了，如果我们最终认为，对这些东西只有主观信念，这些"主观信念"到底正确与否并无事实根据，这样做了之后，说我们还会过得同样好就不对了。

或许我（偶尔）对哲学讨论和政治讨论之间作的类比于讨论有所帮助。我的一位同事，以提倡一切"福利"政府在道德上都应加以禁止这一观点而著称。按他的观点，甚至办学公立制在道德上也是错误的。如果办学公立制和义务教育法（我相信他也会认为这是政府对个人自由的粗暴干涉，同样应予禁止）一起废除，那么较贫困的家庭就无力支付孩子的学费，结果只能选择让孩子成为文盲。但是，他以为，这是私人慈善事业要解决的问题。如果人们连防止大量文盲（或者大量老年人饥饿）等问题产生的慈善心也荡然无存，那就相当可悲了，但政府的福利政策和行为并不因此就是正当的。

依我之见，他的基本前提（例如，财产权的绝对性）是反直觉的，且缺乏充足的论据支持。在他看来，我正陷于漠视个人权利的"家长式"哲学之中。这是一个极大的分歧。这是一个"政治哲学"

的分歧，而不仅仅是“政治上的分歧”。但是许多政治上的分歧包含在政治哲学之中，尽管它们很少像现在这样明显。

在这些分歧中出现了什么情况呢？有时，当双方都很明智地发生争论时，所发生的全部情况常常是一方很敏感地判断并勾画出分歧的由来。通常，当分歧没有达到我们所描述的那样势不两立的程度时，双方也许会或多或少地修改各自的观点。如果实际的一致没有达到，或许对有分歧的双方来说，一个可能的妥协多少还是可接受的。

不幸的是，在不同观点人们之间进行的如此明智的政治争论在当今并不多见。而一经出现，便显得格外令人愉快。在这种讨论中，人们对争论双方的态度十分有趣地包含各种成分。一方面，人们承认并且赞赏某些最重要的理智品质：有开放意识，愿意考虑理由和论据，能接受善意的批评，等等。但是人们所不能同意的基本原则又怎么样呢？假装认为在此理由和观点没有好坏的区别也许是极不诚实的态度。我认为，一个人是否认为团体以同情心待其成员的义务高于财产权不只是一个趣味问题，我的争论对手也作如是观。我们每一个人都认为对方在这个层次上缺少某种敏感和悟性。老实说，我们每个人都有着某种类似轻蔑的成分，不是轻视对方的心灵，因为我们双方都高度尊重对方的心灵；不是轻视对方的人格，因为比起对许多同意我的自由主义政治观点的人，我更尊重我这位同事的真诚、正直和善良，而是轻视对方的情感和判断的某种复合体。

但是，我这就算不诚实了吗？我说我尊重鲍勃·诺齐克的心灵，我确实如此，我说我尊重他的性格，我也没说假话，但是如果我对他的某种情感和判断的复合体（或在棒球场中的某些东西）表示轻蔑，难道不是在对他加以轻蔑（或采取某种类似的态度）吗？

细究下去是极痛苦的事。礼貌通常使我们不能公正地考察我们对我们喜爱而又有分歧的人的真实态度。事实是，任何一个心智

健全的人都不会对某个人(首先不是对他自己)的一切都表示喜爱和尊敬。我们可以对某人从总体上表示喜爱和尊敬,同时又认为他在理智上和道德上有某些缺陷,这两者并不矛盾。正如我们从大的方面喜爱和尊敬自己,同时又承认自己有理智上和道德上(或情感上)的缺陷一样。

我极力主张,在一个有开放意识、尊重理智并有自我批评精神等基本的理智品质的人和一个不具备这些品质的对手之间,在一个具有深刻恰当的实际知识的人和一个腹中空空的人之间;在一个仅仅感情用事、异想天开(人们在讨论政治问题时通常都是这样)之辈和一个深思熟虑之士之间有着根本不同。这种尊敬的轻蔑的矛盾态度是一种诚实的态度:尊敬他人的理智的优点,蔑视其智力或情感的弱点(当然是以自己的眼光来观察的,大凡人类莫不如此)。尊敬的蔑视听上去近乎别扭(特别是当谁把它与轻蔑的尊敬相混时更是如此,但这两者是完全不同的)。如果轻蔑的是作为个人的他人,而不是他人的情感和判断的复合体,那么它也许是别扭的。但是,这是一个远比虚假的相对主义更为诚实的态度。这种虚假的相对主义是,当某人实际地感到一种观点合理,另一种观点不合理的时候,却借口没有拿得出的理由,或者对一个论题的理由没有什么好坏之分,来加以搪塞。

或许,从我们一直在讨论的抽象水平上走出来,再考虑一些相对简单的例子是大有裨益的。试考虑一个大多数普通人在大多数时间准备作的判断:要和平不要战争。(这种判断福科是决不会讨论的,正如斯威夫特决不会描述它们一样,理由是同样的,当这个问题存在时,只有社会庸才,而不是社会贤达才会对此有兴趣。)对此类判断的由来大家都深信不疑。我们对战争的恐怖,对战争给成人和儿童、参战者和非参战者、土地和生灵带来的灾难,仍记忆犹新。尽管这个判断部分地出自自我利益,它并不因此而成为不合

理的，而是恰恰相反。

人人都可以作一个相反的判断，要战争，不要和平。并且，这并不出自合法自卫的理由，侵略和幻想能把人民推向全民族的嗜血好战情绪之中。但是这并没有表明，全部价值判断是不合理的，而仅仅说明，有些价值判断是不合理的，并且表明，当一个人不能摆脱党派偏见或不能批评自己的信仰时，要想说出到底何为合理何为不合理是很困难的。这也是我们在各种认知品质中对公正无私和批判性态度尤为注重的原因。

我们说某些价值判断是合理的和客观的，这并不意味着我们认为对资本主义、民主、社会主义、权利、自治等等的抽象议论总是有意义的。就算我们对某些一般争论想说的都是对的，通常我们也很难将它表述清楚，特别是，如果我们在表达抽象观念方面缺乏训练则更是如此。那些声称没有客观的价值，价值只是相对于环境而言的人类学家的例子正好说明了这种情况。甚至当我们在有效地表达我们想要说的方面如愿以偿时，仍然有一种无理性的巨大力量动摇我们的判断。这里，我的意图不是要否认强力能摧毁我们的判断，自恋能诱使我们的判断出错，我要否认的是，我们在这巨大的力量面前是无能为力、毫无希望的，以至于试图理智而公正地判断是徒劳的（并且，事实上是一种自我欺骗）。说我们可以是合理的不同于说我们可以不犯错误。相反，正如默多克提出的努力寻求一个经过推理的、合理的态度本质上是一件进步的、"可无限完善的事情"。[①]

迄今为止，我所说的一切，一个明智的道德相对论者恐怕也会赞同。一个相对主义者不必费神去掏空所有价值判断的合理性的基础，或者去捍卫福科对历史的描绘，它把历史看作是由毫无合

① 《善的至上性》，第 23 页。

理理由地前后相继的"话语"或"意识形态"所构成的一个不连续的系列。一个更有节制的相对主义者也许很乐意赞同杜威的观点：某些价值是客观相对的，也就是说，某些价值在作出这种价值判断的人的既定的自然环境和历史环境中是合理的。这样一位谦和的相对主义者坚持认为，一切价值的相对性才是至关重要的。他想否认的客观性不是杜威所肯定的客观性（杜威的客观性仅仅是在其实际的生存环境中有根据的判断的客观性），而是柏拉图可能会肯定的客观性，即声称从一个绝对的观点出发，撇开一切场合并且对一切场合有效的、虚构的客观性。

如果我们不满于接受这种有节制的相对主义，如果我们对杜威个人的理论著作感到困惑，我认为，这并非因为我们真正渴求绝对。当我宣称，谋杀和使无辜者受难是错的，我认为，我其实并不关心这个判断对一个具有完全有别于我们的体质和心理的存在物是否有效。如果在半人马座有一种存在物，它不能感到疼痛；不在意个人的生死，很可能我们对谋杀和受难的大惊小怪在他们看来就是杞人忧天。但是恰好是这个有别于我们的生命形式，表明了他们不能理解其中的道德争端。如果我们的客观性是就人类而言的客观性，我们还是有足够的理由称它为客观性。

必须注意的是，杜威的"客观相对主义"原理不能用到纳粹身上（尽管杜威也许会有异议）。我们要说的是，纳粹的目标彻头彻尾地错了；而且，"相对你的利益，这句话是真的，而相对纳粹的利益，这句话是假的。"这一判断就是某种我们颇为反感的"道德相对主义"。客观相对主义对许多道德情形来说似乎是对的，但是对于那些权利和义务清楚明白、界线分明，我们似乎只能在对与错、善与恶之间作出选择的情形而言，就不是这样了。

的确，在某种意义上说，现代的合理性的工具概念本身就是"客观相对主义"的。这个概念的核心是一个极易误解的简单的二分法：目的或目标的选择既不是合理的，也不是不合理的（只要满

足某些最低限度的一致性要求），而手段的选择在其是有效的程度上是合理的。合理性是手段的属性，而非目的的属性。它是完全与效率合而为一的。因此琼斯溺爱香草超过巧克力冰激凌既不是合理的，也不是不合理的。但是假定他的偏好秩序一定，在两个特殊场合选择香草而不是巧克力的行为，对琼斯来说也许是合理的。这个思想，可追溯到休谟的一句名言："理性是且应当是情感的奴隶，"并且还曾对边沁有过深刻影响，它得到了现代社会科学家的广泛认可。现代经济学家的"帕累托最优"概念就是尝试有一个只考虑手段的效率，不包含有关各种经济力量的目标的"价值判断"的经济最优性概念。这个概念对当代颇具吸引力，正是因为人们都假定，手段的选择要服从于理性的批评，而目的的选择则不然。

但是，当我们看到这一理论所依赖的过分简单化的心理学理论时，其全部思想能打动人心的吸引力便丧失殆尽。在边沁的理论框架中，目标，目的，偏好要么被当作固定的个体参数来处理（即，个人的学习活动被勾画为一个学会更好地估量后果和行为的可能后果的过程，是一个学会更有效地达到目的的过程，而不是一个获得新目的的过程），要么（如果目标事实上没有固定）被当作其变化仅仅是没有合理的性质并且是理论家无法加以解释的诸因素的后果的个体参数来对待。

威廉斯[①]已经指出，存在大量的途径，人们可借以对个人的目标，而不只是对其选择来达到目标的手段合理地加以批评。一旦我们超越狭隘的"边沁派"心理学，这种途径便一目了然。

"边沁主义"的上述思想能容纳这样一种情形，在这种情形之中，一个人能够通过合理的批评而同意放弃一个目标（或者不管怎么样，放弃对目标的追求）。这是指这样一种情形，在此情况中，他

---

① 我这里简述了1978年11月在哈佛作的题为"内在的理性和外在的理性"的演讲。

在大大低估达到目标(相对于他的其他目标)的代价方面错误地估计了后果。这就引出了一个不仅同理解力大有关系而且同想象力大有关系的问题:从经验上讲,达到一个目标实际上究竟会是个什么样子?许多人追求他们事实上也许并不真正乐意获得的目标,或者,追求他们不会长久珍爱、或实际上没有他想象的那样珍爱的目标。即使在边沁主义的理论框架范围内,也有可能去完善对合理决定的说明,这些决定是通过考虑错误估计对各种目标的实际存在的感受的可能性而作出的。这开始引进了一种观念:目标本身,而不仅仅是手段,可以被批评为是非理性的。

人们常常一再忽视那些如果想到也会追求的目标。或者,就算他们想到了(或由某人提示出来),他们也许缺乏把达到这些目标时将会实际出现的情况具体化的想象(又是想象!)。特别是当这些目标具有长期的特性(比如提高诗的鉴赏水平)时,更是如此。一个视针戏游戏高于诗歌的人,其实可能并不能想象具有鉴别真正诗歌的细微差别的高度敏感性究竟是怎么一回事。如果他能增长智力,改进想象的话,他就会恍然大悟,他正在犯一桩错误。

一个人所具有的合理地批判自己的目标(以及他人的目标)的能力要依赖其想象力,这和这种能力要依赖其接受真陈述、怀疑假陈述的能力相比,也许占有同等的分量。指出这一点是很有意义的,而且同样有意义的是,一个人的目标也许是精神或性格的持久特征,而不是一件事情或一个事件。

除了错误估计一个人的目标的真实经验之外,或者,除了错误估计可能使人犯选择目标方面的错误的另一些可供选择的目标的真实经验意义之外,还有其他一些方面。威廉斯指出,重提一个可追溯到亚里士多德的观点,通常一个目标是总体性的(如过一个愉快的夜晚),问题不全是为"目标"找到一个手段,而是找到一个包括达到目标的可接受的详尽计划的全盘活动方案(例如"去看电影","在家读书")。一个人能否对其目标构想出一个有创造性而又

新奇的计划，抑或只想到一个平庸而乏味的计划，这还是依赖于想象而不仅仅依赖于对命题的理解力。

正如威廉斯所指出的，问题在于，即使某个人用一个对所有这些事情都一视同仁的解释来取代狭隘的边沁主义心理学，他还是有某种相对主义之嫌。威廉斯所举的例子是，假定有一个年轻人，其父希望他从军，以戎马为生。老人搬出家庭传统（祖上数代男性皆为军官）和爱国主义，企图说服他，结果枉费口舌。甚至当这个年轻人尽其所能把作军官的前景向自己描绘得活龙活现时，这一目标对他仍无任何吸引力。这不是由于理解力和想象力的无能，而只是因为这不是他的目标。

甚至纳粹的情形也与之相似。假定纳粹赢得第二次世界大战的胜利，这样我们就不能把德国的失败作为不当纳粹分子的实际理由。也许某些纳粹分子完全缺乏有关纳粹主义的实际后果（所招致的灾难和痛苦等）的知识。也许，某些纳粹分子如果具有理解这些后果、或更生动地理解另一种生活、一个好人的生活的理解力和想象力的话，他们就不会当纳粹分子了。但是毫无疑问，许多纳粹分子依然会当纳粹分子，因为他们并不关心他们的行为带来的痛苦，还因为无论他们在想象中把那另一种生活弄得多么活龙活现，也不会有威廉斯的故事中戎马生涯对青年人所达到的效果。在他们那里并没有一个可以诉诸的目的；既无现实目的，也无潜在目的，也就是说，没有一个如果他们多一些理解力、多一些想象力就会加以实现的目的。即使没有"边沁主义心理学"，我们还是一再面对道德相对主义的问题。

我们来考虑一个比纳粹要温和一些的例子。想象一个因某种原因对艺术、科学（能帮助他们耕作的一些产品除外）、宗教，一句话，对一切精神或文化的东西全无兴趣的农家社会（我并不是要暗示实际的农家社会是或者曾是如此）。这些人不必被想象成坏人。只要你愿意，你尽可以想象他们互帮互助，和平友善，通情达理。我

希望读者想象的是，他们的兴趣只限于相当贫乏的程度。只求温饱，不计其他。能在辛劳之余和人对酌几盅便是极大的享受。一句话，想象他们是以更接近于动物性的存在方式而生活，而不希望以其他存在方式生活。

这种人不是不道德的。他们的生活方式全无禁忌。但是(除非我们沉湎于道德相对主义之中)，我们的自然反应却是，他们的生活方式在某种意义上讲是低贱的，在亚里士多德所说的“高贵”中完全没有这种情况。他们在过着猪一般的生活——或许是讨人喜爱的猪，但依然是猪，而猪的生活终究不是适合于人的生活。

同时——这是一个难点——我们也不愿说这种猪-人在任何意义上是不合理的。说他们是不合理的，这也许是我们对边沁的“合理”和“不合理”的用法的长期文化适应的结果。但是，尽管这样，它成了我们现在的内在气质。我们要说，这种猪-人的生活是不如他们应有的那般好，但他们并不是不合理的。

我们无意声称，人们的生活过得是好是差，这只是一个趣味问题。我们不明白，我们何以能说选择一个更好的生活是合理的，而选择一个较差的生活是不合理的。然而这样一来好像恰恰是说：“一切都是相对的。”我们赖以立足的基础便倒塌了。

由威廉斯所提出的对边沁主义心理学的某些修正或许对此难题有所帮助。我们假定猪-人出生时有正常人的潜力(如果他们没有这种潜力，他们的生活就不会“不如他们应有的那么好”，而且，我们顶多只能加以怜悯，对他们加以蔑视则是不正当的)，然后，他们也许被引导欣赏生活中的艺术、科学和精神的各个方面；也可以说，被引导过一种更适合于人的生活。如果他们之中有谁真的这么做了，毫无疑问，他会喜欢这种生活(尽管这种生活不很好笑)，而不要他现在的生活。曾过着猪一般生活的人们，当他们终于过上更适合于人的生活时，会感到羞愧。而过着更适合于人的生活的人，当他沦为猪类之后，不会为他过去的生活感到羞耻，这些事实使得

人们有理由认为，猪-人正犯着某种如威廉斯所讨论的认知不足的错误；有理由相信，他们忽视了另一些可选择的目标，并且当然有理由相信，在他们头脑里，这些别的可选择的目标一经实现会是个什么样子还是一抹黑。一句话，其实并不能说他们选择了较差的生活；因为他从没有一个关于较好生活的合适观念。

虽然我们或许可以由此指出，这些人的生活可以进行理性的批判，但怎样把这套说法贯彻到纳粹身上还不明显（人们可能会同义反复地说，谁没有实际地选择较好的生活谁就没有"恰当地设想过这种生活"。但这种花招恐怕显然是无济于事的）。即使在猪-人的例子中，如果他们是一些意识形态上盲从的猪-人（与单纯的猪-人相反），那么我们羞愧的观点就无法成立了。在此情形中，我们无法求助于任何属于他们的，哪怕是潜在的目的。

对猪-人在理性方面的缺陷，我们不愿过分吹毛求疵（除非我们能指出他们所没有达到的目的，至少是潜在的目的），这是我们文化中的理性概念近年来变化的结果，这一点是不难证实的。因为不论古代哲学家还是中世纪哲学家，当他们听到这样一种说法——如果A是比B要好一些的生活，那么这一事实就是选择A而不是B的最有可能的理由——时，都不会感到奇怪。目的的善能够使选择该目的的行为变得合理，但是我们已经丧失了弄清何以能如此的能力。

当然，在很大程度上，我们可以用我们不承认善是客观的东西这一事实来解释上述无能。但是现在我们面对一个循环，或者说两条曲线。一个是现代的循环：合理性的工具主义观点支持这样一种说法，一个目的的善并不能特别地使得不选择那个目的或选择一个坏透了的目的的行为变得不合理。这又支持这样一种说法，即善和恶都不是客观的。这种说法又支持这样的论断，即工具主义的合理性观念是唯一明智的观念。一种是传统弧线：理性是一种依据目的的善（而不是"激情"，它力图根据趣味来指派目的；或"倾向"）。

这种主张支持选择善是合理的这一观点。这个观点又支持了善性和恶性是客观的这一说法。很明显，我们不能简单地回到古代或中世纪的世界观上去，而不管保守主义会希望什么。但是，难道我们除了边沁的循环，便别无选择了吗？

第八章　科学对现代各种合理性观念的影响

我们关于合理性问题的讨论，持续了数十年而终无结论，这或许是因为这个讨论从来就假定了合理性之于善的优先性。问题始终在于，我们是否可以说选择了一个坏的目的，即可被称为不合理的，似乎善是被告而合理性是法官。特别是当某人对合理性的假定在很大程度上只是未经检验的文化偏见和神话的拼凑，坚持合理性高于善这种立场就是预先断定价值判定的性质。在本书余下的篇章里，我打算转换比较的角度，不问合理的怎样成为善的，却问是合理的何以就是善的。询问价值合理性本身所具有的性质，将有助于我们澄清合理性本身的性质，澄清我们对合理性所易于作出的种种假设，并且使我们得以明白，我们思考前一个问题的方式究竟错在何处。

我们可回忆一下，当马克斯·韦伯提出事实-价值的现代区分时，他用来反对价值判断客观性的论据是：不可能制定一个使所有可能的理性的人都满意的价值判断的真理[①]。从一开始，使人们对价值判断多少有点持怀疑态度的正是合理证明的不可能性，或据

① 特别参见"社会科学知识的客观性"，载《社会知识及社会政治文库》，第19卷（1904年），第24—81页；"社会科学和经济科学中的价值分离概念的意义"，载《逻各斯》，第7卷（1917年），第49—88页以及学术报告，"作为职业的科学"，1919年，这三篇文章都重印于《社会科学方法论》，格伦科，伊利诺伊，1949年。

说存在着不可能性。价值接受合理性的审判久矣；[①]这是"实证科学"的结果。实证科学据说能够建立在使所有合理的人都心满意足的基础之上。评价合理性的一个理由是显而易见的。科学的合理性不可否认地有助于我们达到各种实践目标。尽管没有多少有教养的人士会同意我们只应为技术的成功而追求科学这一观点，但毫无疑问，说科学在技术上的成功是压倒一切的，这决不是夸大其词。我们生活在一系列显然是不停的技术革命——"工业革命"、"电子革命"等等之中，这些技术革命不断提醒我们，科学对于塑造我们的生活是一个多么巨大的力量。甚至在工业革命之前，牛顿物理学的空前成功便在无数人的心灵中留下了深刻印象。例如，17世纪刚开始讨论"进步"概念时，相信社会是不断前进的进步论者用"牛顿比亚里士多德更博学"来证明自己的观点。没有人能令人信服地证明莎士比亚是比任何一位古代悲剧作家更优秀的剧作

① K·O·阿佩尔在"解释学和伦理学的共同预设：超越科学和技术的合理性类型"（载《现象学研究》，第9卷，1979年）一文中，对韦伯作了和我一样的理解。阿佩尔这样写道（第36页）：

> 但是，马克斯·韦伯对我提出的关于超越价值分离和技术的合理性的可能类型的问题，也持严厉的否定态度。我以为这种回答已成为当代西方意识形态体系的典范。韦伯将方法论的理解范围局限在他以"有目的的合理行为"的"有目的的-合理的理解"的"理想类型"这一概念为中心的"价值分离"的理解之中。而"有目的的-合理的行为"，或许也可称为"工具性行为"，在这些行为是成功的那些情况下，它们可以被分解为或重构为建立在从规律科学的如果-那么规则到技术指令的如果-那么规则的成功转换的基础之上的。因此，马克斯·韦伯把方法论的理解局限在设法掌握超越人类行为的（价值分离的）技术的手段-工具-合理性之中。这实际上构成了韦伯的限制性的合理性范型的工具合理性的思想。
>
> 但也必须注意到，在社会学中，对有目的的-合理的理解来说，并不是非得满足一个最高的要求不可，即确定行动者成功地把规律的规则转换为关于手段-目的关系的技术准则。为了能借助那种工具合理性来理解其行为，只须确定，他在其目的，以及其有关有助于达到目的的手段、方式或策略方面的信念的先决条件下的行为是合理的，便足够了。因此，经验-解释学的理解活动，就在于假设性地找出并证实行动者的这些目标、意向和手段-信念，根据这些目标-意向和手段-信念，他的行动可以在技术的手段-目的合理性的意义上被理解为合理的。

家，或是一位比荷马更出色的诗人。但是不可否认，和古代科学家亚里士多德相比，作为科学家的牛顿在知识上取得了切实的、不可否认的进展。

尽管百科全书派和其他一些人很快将进步概念从科学领域推广到政治制度和道德领域，这种概括在 19 世纪是理所当然的事，但是在 20 世纪则受到了怀疑。

孔德建立了实证主义哲学，为科学的成功歌功颂德。在孔德那里，历史是一个成功的故事。我们从原始神话开始，这些神话随后变得精致、纯净，直到高级的宗教出现；这些宗教又让位给柏拉图或康德的形而上学的理论。最后，到孔德的时代，它们又不得不让位给"实证科学"。很显然，没有人会怀疑谁是这个成功故事的主角。这个主角就是科学。如果当初科学留在少数人心目中最深的印象是其惊天动地的理智上的成功，那么毫无疑问，对直到今天的大多数人来说，科学给他们的最深印象一直是其压倒一切的物质上和技术上的成功。关于这一点，哪怕科学威胁了我们的生活，我们也是印象深刻的。因此，对价值判断具有认知性质持怀疑态度的原因之一，就是价值判断不能"被科学方法所证实"，正如人们一再指出的那样。同样，在伦理学问题上不可能得到普遍的、哪怕是大多数人的一致意见，这也是一个事实。我们已经看到，这一点在福科的讨论中起着重要作用。堕胎是对是错，同性恋是对是错这类问题，与那种一人回答便人人心满意足的问题似有不同；而科学理论的正确性能够被证明得让人人心服口服，这是绝大多数人坚信不移的。他们认为，科学本身的合理性在于，或者部分地在于这样一些事实：科学的预定至少能够被公开地加以证实；对科学预言的结果是否能获得，对该理论所预言的现象是否确实出现，没有异议。当然，这里有一循环论证的危险：如果我们把理性的程序等同于导致多数人赞同的结论的程序，那么韦伯关于在伦理问题上，一个人即使正确，也不可能获得所有理性的人的赞同的观点就会意味着，

一个人不能博得所有使用一种据称保证赢得半数人或绝大多数人赞同的方法的那些人的赞同。这即是说，我们用以确定价值判断不能作出令所有理性的人们满意的证实的方式，仅仅是根据观察到它们不能作出令总人口的绝大多数满意的证实。这毕竟并不等于说我们拥有对合理性的检验手段。韦伯的理论包含的意思是，我们多少得先清点一下理性的人的人数，然后再看能否使他们一致赞同某些价值判断的真假。但事实上不可能真有此事发生。韦伯的所有例子(中国的达官贵人等等)所真正说明的是，价值判断不能按照令所有有教养的人或理智的人(这与所有理性的人无疑是一回事)都满意的方式得到证实。韦伯的论证是一种经过伪装的多数主义的论证。他正在求助于这样一个事实：我们能够在实证科学方面博得有教养者的一致赞同，而在伦理价值方面则办不到。将这种观点和亚里士多德的观点作一比较是很有意思的。亚里士多德说，在伦理学问题中，我们当然应该永远争取得到大多数人的赞同，然而我们知道，从实际情况来讲，我们又通常不能如意。有时，我们只能相信智慧；而且我们必须依靠自己的判断来决定一个人的智慧与否。尽管这在今天听起来像精英统治论者的论调。

当然，这并不是说，一个人当真能够使一个被人任意接受的科学理论的真理获得绝大多数人的赞同。事实是，大多数人都令人遗憾地对科学一无所知，而许多理论，尤其是精确科学中的理论，要理解它们所必须具备的数学知识之多，使得大多数人甚至没有能力去理解它们。尽管大多数人对此未加否认，但这似乎并未影响他们的观点。因为，根据淡化了的操作主义的观点(它似乎已成为大多数科学家的工作哲学)，科学理论的内容在于可检验的结论，而这些结论可由以下形式的陈述所表示：如果我们实施如此这般动作，则我们将得到如此这般的可观察的结果。这种形式的陈述，如果是真的话，可以通过足够频繁地重复适当的实验来说明其真。不错，这种解释有许多困难，比如，实验的设计、实施和评估比门外

汉想象的要难得多。但是，毫无疑问，在精密科学中，对某些理论的实验的恰当性达到广泛的赞同，事实上从来就是可能的。门外汉们对这些理论的接受也许是出于对专家们的尊重，但是至少专家们似乎是意见一致的。

当然，从理论上讲，工具主义本身并不构造并且也不会自行构造一个可靠的合理性观念。科学结果无疑有巨大的实践价值。但是，正如我们所说的，没有一个有教养之士会说，科学只是出于其实际运用的目的才具有价值。就算科学只是出于实际运用的目的而具有价值，为什么合理性只由其运用之故而有价值呢？不错，拥有一个能帮助我们选择有效的手段以求达到我们不同的目的的工具是有价值的，但是知道我们应该选择什么样的目标也是有价值的。如果“合理性的证实”依其定义而被限制在手段-目的的联系的确立之上，那么价值判断的真理不可能被合理地证明就不足为奇了。但是，为什么我们应首先有这样狭窄的合理性概念呢？①

---

① 阿佩尔把这个狭窄的合理性概念归之于韦伯，他写道：（见上述引文，第 37 页）

> 韦伯的“理解”的方法论这个问题和他的多少有些隐含的历史哲学是完全一致的。因为，在他自己对西方文明史的重构的上下文中，他从一个颇具启发性的假设出发，即，历史的这一部分至少应被看作是合理化的不断进步，同时也是一个不断幻灭的过程。用他自己的话说，就是一个不断“解惑”的过程。对“合理化”概念，他理解为在科学、技术进步的不断影响之下的这样一个进步过程，即把手段-目的-合理性的运用范围扩展到社会-文化系统的各个方面，特别是经济和阶层统治的领域。而对“幻灭”和“解惑”的过程，韦伯的一个理解则是指一个为公众所接受的宗教和哲学的价值秩序或世界观的解体。就他个人的世界观而言，他准备从这个过程中得出一些实用的结论，因为他认为一个严谨而真诚的思想家必须接受下列见解：“合理化”意义上的进步的补偿在于，放弃对最终价值或规范进行合理估价的想法，以便在面临关于最终价值或规范的多元论，或者如韦伯所说的“多神论”的情况下求助于最终出自良心的前合理决定。

多数主义也是站不住脚的。一个人视为真的东西能获得赞同当然不错。避免和其同伴发生冲突也总是一桩好事。但是千百年来人们在生活中一直知道这样一条并不令人愉快的道理：在某些事情上，一个人不得不依靠他自己的判断，即使它和大多数人的判断相左。许多人为相信自己不同于多数人的判断而自豪。在某些问题上（伦理学是其中之一），所要权衡的事复杂而且含混，以致我们不能指望任何科学证明或科学定义，而只能依靠感觉和判断，这个观点由来已久。事情大概真是这样的，合理性的最高表现之一，恰恰是在人们不指望令多数人满意地证明一件事情的地方进行正确判断的能力。某些事情不可能人人满意地加以证明这一事实居然成为关于这类事情的信念的不合理性的证据，这是不可思议的。[1]

尽管这些观点在理智上是经不起推敲的，但是工具主义和多数主义特别受当代精神的青睐似乎也是实情。当代精神偏爱的是可加以证明的成功；当代精神所讨厌的恰恰是判断力和智慧这些概念。我不是一个社会学家，我也无意探究何以工业社会（在其资本主义和社会主义的形式中）应该如此热衷工具性的成功并且多数人同意这些主题。但是无疑，这种社会学的事实与合理性等同于科学的合理性这一观念、与科学的合理性本身的基础在于表明和使绝大多数人（潜在地）满意有工具性联系这一观念的日益增长的重要性分不开。

我们适才讨论的对合理性的看法，把合理性看作是导致对有效的手段-工具联系的发现和对这些联系的“公开”确定的种种方法（不管其本质是什么；这些方法的本质通常仍然是含糊不清的）。如果这种看法在理智上并不像初看之下那么站不住脚的话，使它变得可尊敬一些的哲学努力却一直并不少见。努力之一来自洛克、

① 的确，我们在第五章已经看到，有关合理的可接受性的一致论是自相矛盾的。

贝克莱和休谟的经验主义。及至穆勒的时代，这种经验主义凝固为哲学家所称的现象主义，这种学说认为，我们能够切实谈论的只有感觉。根据这种学说，甚至日常对象，例如桌椅等，真正说来也只是人类现实和可能的感觉中的客观规则性的集合。如穆勒所说的，物理对象只是"感觉的恒久可能性"。这同一种观点的另一种说法是，所有看似对物理世界或诸如此类东西的谈论，实际上只是经过好几重推论而进行的关于感觉的谈论。

这种观点在其持有者眼中的长处在于，它不仅能使他们明确地说出科学的内容是什么，而且能说出所有者认知意义的任何谈论的内容是什么。任何科学的理论其实只是陈述大量下述形式的事实的"经济"方式：如果你实施如此这般的行为，你将得到如此这般的经验。持这种观点的人不必捍卫这样一种靠不住的观点：科学家只热衷于应用，或只热衷于达到实用目标，而对为知识而知识则毫无兴趣。现象主义者不必否认，我们想知道黑洞的本质，也不必否认，我们想知道是否有一个大爆炸，以及人类的真正起源是什么等问题。我们确实想知道这些事情，而且不仅仅因为知道这些能使我们制造更好的机器。但是知道这些事情（尽管这要用精确的哲学分析来显示）恰恰就是知道大量的以如果你实施如此这般的行为，你将有如此这般的经验这种形式出现的事实。无论我们热衷于这些东西的动机如何，所有的事实最终都是工具性的。同时，对这样一种论断，即，称某物是好的，等于作出了如果你实施如此这般的行为，你将有如此这般的经验这种形式的预言，我们似乎还无法加以理解。因为根据他们的观点，涉及任何事情的好坏性质的陈述都没有认知的意义。用 20 世纪逻辑经验主义的话说，这类陈述纯属"情感"的陈述。然而现象主义在两个方面失了足。首先，关于物质对象的陈述可翻译成关于实际感觉和可能感觉的陈述这一论断事实上似乎是不能成立的。从 30 年代卡尔纳普和维也纳学派的工作开始的，对这一论断的详尽的逻辑考查，使现象主义者自己也相信

上述说法是无根基的。作为总体的科学理论无疑会引导我们期待这样一种情形发生：一旦我们实施某些行为，我们将有某些经验。但是，科学陈述能一一对应地翻译成关于一旦我们实施某些行为、我们所将得到的经验的陈述，这一思想现已被当作一种不可接受的还原论而放弃了。其次，感觉必然是私人对象，尽管我们也许能够在实践中通过干脆问他们提问来决定某人是否有某种感觉。然而，一旦某人提出："你怎么知道那人像你一样把同一些感觉与他的描述联系起来呢？"我们便立即陷人一种认识论的窘境之中。如果科学的内容在于预测，一旦理性的存在物实施某些行为，这个理性的存在物所要得到的是什么感觉，那么要知道这种内容本身意味着什么，我们必须要能够回答外星人（如果我们有幸遇到的话）是否具有和我们相同的感觉。为此，诸如卡尔纳普和波普尔这样的哲学家认为，科学的观察性预测应以下述方式表述：如果有人实施如此这般的行为，则将发生如此这般可公开观察的事件。在这里，实施的行为和应期待的可观察事件都必须借"公开的"对象（如代表读数）来描述，而不能用感觉这种私人对象来描述。

总结一下，老派的经验主义或现象主义似乎在塞给我们一个令人满意的关于认知的意义标准：如果一陈述可翻译成关于感觉的陈述，则它便是有认知意义的。但是事实证明，要么"翻译"概念相当含糊，要么科学陈述本身不能满足这个认知意义标准。根据事实陈述可翻译成关于感觉的陈述、而价值陈述不能作如此翻译来为两者划一条泾渭分明的界限，其麻烦之处在于，所谓的事实陈述的可翻译性至今未能并且不可能得到证明。经验主义的还原论在事实和评价之间作了明确的划分，然而这是以对事实的完全歪曲的图景为代价的。

但是，我们的初衷是考察"合理的为什么就是好的"这一问题的答案。我们考察的第一个答案是，合理性能使我们发现可靠的手段/目的的联系。这一说法由于太狭窄而被我们拒绝了。我们之所

以提到现象主义，是因为，如果现象主义是真的，那么在热衷于科学理论的两种动机之间的明显冲突，即，为了工具性后果而热衷于科学理论和为了弄清科学理论对于自然过程到底告诉了我们一些什么而热衷于科学理论这两种动机之间的冲突就会消除。在工具兴趣和纯理论兴趣之间的冲突在某种意义上可以巧妙地加以消解。当然，在这些兴趣之间还会有某种区别。但是，从最终的逻辑分析来看，哪怕再纯粹的理论兴趣，也是对一直展现着工具本性的事实的兴趣，一切名副其实的知识向来都可以被表明为手段/目的关联的知识。只不过，当我们想用手段/目的关联来达到某些目的而对它产生兴趣时，我们就称这种兴趣为"实践的"兴趣；当我们纯粹出自好奇心而热衷于搞清手段/目的关联时，我们便称之为理论的兴趣，那种将所有科学陈述都归结为一旦你实施行为A、你便将得到结果B这种形式的陈述的企图，如我们已经提出的，业已破产。

作为科学家的实践的一种概括，卡尔纳普和波普尔关于科学的观察陈述是用物理事物的语言而不是用感觉语言来表达这一论点显然是正确的。但是，假若把这一点上升为一个认识论上的绝对，则它的影响就更为巨大。一方面，如果根本就不允许观察陈述谈论感觉，那么内省作为一种科学的观察方式就被排除了。尽管许多心理学家会同意这种做法，但实际上，还是有一种心理学家不为哲学和心理学的教条所束缚，继续从事至少是部分地包含着对内省报道的信赖的实验。事实上，卡尔纳普也许并不像行为主义心理学家那样武断地禁止这类实验，只要感觉不是当作观察报道而是行为材料，这里"行为"指的是作出这些言语报道本身的行为，卡尔纳普也不会反对使用感觉报道。但是，将接受一个内省报道当作"从一个言语行为出发"的推论这究竟是什么意思，还不甚明了。即使是在关于物理对象的报道（例如，我面前有一张桌子）而不是关于感觉的报道的情况下，我们通常也不接受这种报道，除非我们拥

有这样一种理论，据此，一个人能够观察到他们报道的事实。从这个意义上讲，观察本身应该具有理论上的可解释性便是我们对世界图景的融贯性的全部要求的一部分；如果有人宣称他靠特异功能观察到在某处有一张桌子，我们不会接受这个“观察报道”，因为它与我们理论的总体内容不融贯。在此意义上，每一观察报道都有某些“推论”的成分。另一方面，当一位医生接受病人感到疼痛的自述时，这医生根据什么“科学理论”从病人的言语报告中推出该病人感到疼痛的结论是难以知道的；如果把认为人们完全能够说出他们是否感到疼痛这个一般假设算作一种理论，那么人们完全能够说出他们面前是否有一张桌子这个一般假设也算作一个理论。但是我们很难看出，在接受某人关于有一张桌子在他们面前的说法和接受某人关于他们感到疼痛的说法之间，有什么根本的方法论上的差别。

波普尔和卡尔纳普也许会回答说，其认识论上的差别在于，前一种陈述是可公开检验的，后一种则不行；但是他们都夸大了物理对象的观察报道永远可以公开检验的程度。这样的报告，许多都是要借助于需长期训练才能使用的工具完成的。（众所周知，学会从一个高倍显微镜下“看”东西需要大量专业化的训练和技巧，并且不是人人都能获得这种技巧的。）接受这个认识论的教条的实际含义，就是使合理的信念是能够公开检验的成为合理性的定义的一部分。让它成为合理性定义的一部分是很容易的，那样你就无需为此观点提供任何论证了。或许这个证明说到底是这样的：凡是不能公开检验的，就会造成分歧，而在有不可调和的分歧的问题上，无所谓对错。但是这样说也许假定了我们所说的多数主义，即，能够合理地加以证明的，即是能令绝大多数人满意地予以证明的。这种思想已经成为合理性概念本身的组成部分了。

逻辑经验主义是对工具主义和多数主义这种主导性的文化倾向的成熟的表达形式，这一点在其后期的发展中变得愈加明显，至

少我以为如此。尽管逻辑经验主义早在1936年就已放弃了现象主义，但是在随后的20年里，逻辑经验主义的科学哲学家们热衷于讨论科学的目标，热衷于把科学的目标当作是预言(带有我随后要讨论的某些附加条件和限制)，这种现象一直延续到逻辑经验主义作为一个引人注目的哲学思想消失解体为止。科学的目标就是预言，这一思想从孔德开始就是实证主义的基本信条。正如我们所看到的，这一思想在现象主义流行的时代，至少有某些严肃的哲学理论基础。因为在那时，一个人可以辩解说，所有有认知意义的陈述都是伪装的预言，或者是伪装的预言的无穷集。在现象主义消失之后，这个学说的再现恰似一个病人在治疗过程中，当"保护层"被剥去之后，其联想的"原始"材料的浮现。认为科学的目标是成功的预言(或成功的预言外加干脆称之为"简单性"的某些东西)，这种观点和认为我们只是出于实用目的而追求科学仅仅一步之遥。没有哪位哲学家愿意坚持这种危险的观点。不错，那些为纯工具主义科学观而辩护的哲学家会坚持这种观点，但他们这样做并非因为他们是实践的崇拜者，或者是不配为其他目的而欣赏抽象的科学知识之美的目光短浅之辈，而是因为，他们认为，只要将那些有"认知意义"的东西看作是对形成预言有价值的东西，他们便能一劳永逸地排除各种形式的蒙昧主义和形而上学。对这些哲学家来说，"形而上学"只是各种先验的思辨的别名。他们唯恐避之不及的，正是这种宗教和"形而上学"的思辨(他们意义上的"形而上学")。

科学的成功把哲学家们催眠到如此程度，以致认为，在我们愿意称之为科学的东西之外，根本无法设想知识和理性的可能性。故我这里想说的是，在有这样一种哲学思潮的文化中，考虑到科学在一般文化中的崇高声望，考虑到宗教、绝对伦理学和先验形而上学的名声日渐低落，出现上述情况便是可想而知的事情。我还想说，科学在一般文化中的崇高声望要极大地归功于科学的工具性的巨大成功，以及科学似乎摆脱了我们在宗教、伦理学和形而上学问题

上所能见到的无休无止的、不可解决的争端这一事实。

但是，既然那些将我们文化中的工具主义倾向合理化了的专业哲学家本身不是情趣低下、只重实际的人，因此，他们自己觉得应该放宽对科学目标的理解，以便更明确地为单单是成功的预言之外的其他目标留下余地，就不足为奇了。所以我们发现了逻辑经验主义者们在40年代和50年代所列举的其他目标：发现规律，追述（对过去的，而不是未来的事件的预言）以及发现各种“说明”。所有这一切在他们看来只是根据规律作出预言和追述。

这里的情形是饶有兴味的。为了阐明科学之热衷于发现自然规律，纯由自身之故，而非仅仅为了获得这些规律所导致的预言，这些科学家列了一份科学目标一览表，以取代“科学的目标是成功的预言”这一简单公式。这份一览表可无限延长。结果，自然规律不仅包括严格意义上的自然规律，即那些在物理学上不可能出错的陈述；也包括所谓进化论的规律，这个规律其实是对某种趋势的描述（这种总趋势，由于理智生活行为的缘故，也许会在某个时刻不再保持）；甚至还包括有关个体群和个体有机体的纯粹偶然的特征的陈述。我们当然可以说，科学家的工作是试图发现“自然规律”，包括长期适用、且有较强解释意义的、从物理学上讲是偶然的那些一般概括，诸如，进化论所赖以建立的一般概括，或经济学所赖以建立的一般概括。我们还可以说，科学家试图发现有关群体和个体有机体的特征的有意义的真理，并试图将所有这些编织到一个演绎或归纳的结构之中。但是列这份特殊一览表的理由何在？

理由在于，人们想使这份一览表宽泛到足够容纳物理学家们所追求的各类真理，同时又狭窄到不含任何可反对的（无认知意义的）材料。传统的认知意义的标准，例如“一句子是有意义的，当且仅当它可能被证实或证伪”，对这种标准的要求已被一个关于各种陈述类型的一览表所取代，从而一陈述如果是该陈述类型之一员，就可以被接受，否则则被拒绝。但是哲学家究竟怎么会采取这个步

骤的呢？即使我们称作"科学"的所有部门中的所有陈述果真都是这些类型的陈述(历史的说明当真只是把各种追述归在"规律"之下吗?对这一点我们还不甚明了),我们是否可以由此说,这些类型的陈述的证实,并且只有这些类型的陈述的证实是理性自身的目标,而不仅仅是作为理性的特殊运用的科学的目标呢？当然,我们的答案是,这些哲学家并不认真地怀疑,科学穷尽了理性的全部外延。但他们为何对此不加怀疑呢？因为在他们看来,对立不存在于科学(也就是基本上采用经验科学方法和数学方法所从事的认识)和非形式理性(这种理性是采用那些迎合科学兴趣之外的兴趣的方法而活动的,但未必就不能拥有些真正的标准)之间,相反,对立出现于由科学方法所获得的知识和自命由神启、或某种荒唐的先验才能得到的伪知识之间。理性必须与科学具有同样外延的观点将这些哲学家推到某种奇特的窘境之中。既然他们无意否认(比方说)历史知识的存在,他们也就承认了历史是一门科学,甚至承认了,一个历史学家真正所做的,是把对过去的个别陈述归于规律之下,而这种对历史的断言从表面上看却似乎是错误的。

也许,逻辑经验主义思潮到20世纪50年代开始衰退是在情理之中的。我们一直是仅从一个问题的角度来看待这一思潮,而逻辑经验主义者具有大量不同的哲学兴趣,在许多方面都作出了较有价值的贡献,然而,对于我们一直在追问的"合理性"好在哪里这一问题,逻辑经验主义运动代表了对以下观点的一个经过推理的哲学辩护,认为此问题的答案,且唯一的答案是,合理性有益于发现手段/目的的关联。现象主义的哲学理论为逻辑经验主义者们提供了一种有趣的哲学辩护。一旦现象主义遭到摒弃,并且,对这种主张的哲学辩护被一个空洞的主张所取代,而且更进一步,一旦这种空洞的主张通过引入例外、修改而被炮制得更加"合理"一些,这个运动的全部锐力就随之消失殆尽。将理性的自身目标看作是预言、追述和自然规律的发现,以及所有这一切的系统化,并且,将

这些看作是理性目标的全部内容，这种观点的困难之处在于，我们恰恰没有任何理由去相信它。这并非意味着我有理由相信它错了，如果一个自然规律的概念被放宽到将个体有机体特征方面的陈述的发现也包括在自然规律的发现之中，并且，这个特征概念也非常宽泛（即非常含糊），以致某一科学家妒忌同事的名气这一陈述也可称为一个特征方面的陈述，并且，如果说这位科学家开了某个玩笑是因为妒忌他同事的名气这一陈述也是“把一个特殊事件归并于规律之下”，那么，事情就变成了，一个人所说的一切，都可以被诠释成要么是在陈述一般规律，要么在把一些描述归并在一般规律之下，甚至连说某人是道德善良的人也可以被看作是把一种“特征”归之于某人。这当然不行。试图根据这类目标一览表来详尽地说明一般认知探究的目的，这样做的困难在于，该一览表本身不得不逐条加以说明。如果这个一览表中的诸条只按其字面意义来说明，那么在此表中的陈述种类就不会将科学家们所热衷发现的全部陈述种类都包括进去。特别是当“科学家”也包含那些历史学家、精神病医生和社会学家时，情况肯定如此。而一旦我们对一览表中的诸条规定得如此宽泛，以致毫不费事就能说明属于该表中的、由历史学家所作出的陈述，以及日常心理学的语言表达的描述性的陈述，则此表就变得毫无价值。无论如何，在对为什么这些类且唯有这些类陈述能合理地证实缺乏一个认识论说明的情况下，这样的一个一览表也许是对理性探究极限的一个纯粹的假定。而一个假定，不管是以一览表的形式，还是以其他的什么形式，都不可能具有逻辑经验主义希望“认知意义标准”所具有的排斥力。

“方法”崇拜

既然对“合理的为什么是好的？”这一问题不能简单地回答说，合理性能使一个人获得实用目标，也不能简单地说，合理性能使一

个人发现手段/目的的关联，我们就应该考虑在不同时代有相当号召力的另一种可能的回答：许多哲学家一直相信，科学活动是通过遵循一个独特的方法进行的。如果事实上真有这样一种方法，借助于它，一个人可保证发现真理；如果其他方法都没有发现真理的真正机会，并且，如果正是科学且唯有科学对这种方法的始终如一的运用，才能说明科学的非凡成功和非科学领域的无休止争论，那么，合理性（如果有这样一种东西）也许应该等同于这种方法的拥有和运用。对"合理的为什么是好的"这一问题的回答因此就是，合理的即是好的，因为一个人如果是合理的，他就能发现他所感兴趣的任何东西的真理。反之，如果一个人不是合理的，他便没有真正的机会发现真理，除非是由于运气。和工具主义的观点一样，这种观点经历了兴起、停滞和衰落的哲学历程。从19世纪40年代穆勒的《逻辑》一书发表，到卡尔纳普的《或然性的逻辑基础》问世，权威的科学哲学家一直相信，经验科学的基础是由某种类似于形式方法（"归纳逻辑"）的东西所构成，他们并且相信，不懈的工作或许导致了该方法的清晰阐明，导致了一个堪与弗雷格在1879年出版的《概念演算》中开端的演绎逻辑的形式化相媲美的归纳逻辑的形式化。如果这种方法果真被发现过，即使它本身并不能证明这种方法穷尽了合理性的外延，但对于那些认为存在着可以用其他的方法来证明或表明是可合理接受的真理的人，却提出了一个证明自己观点的棘手任务。

根据影响甚广的所谓贝耶斯学派的观点，哲学家们一直试图形式化的这种归纳方法的总的特征是：我们假定科学语言已被形式化了；科学家们拥有一定数量的可由这种形式化语言中的"观察语句"表达的可靠的观察。我们还假定，各种正在考察的假设都是由这种形式化语言的公式来表达的。归纳逻辑的问题被看作是定义一个"确证函数"的问题。这个函数是概率函数，它将确定每个假说相对于观察证据的数学概率，或者换种说法，确定该证据提供给

每一可选择的假设的“支持度”。人们通常假定，如果每一可选择的假设为真，他们知道可能获得的既定证据的概率，这就是所谓的“正向概率”，即假设既定情况下的概率。我们希望测算的是所谓“反向概率”，即，证据既定的情况下假设的概率。贝耶斯定理将这个反向概率作为正向概率和其他一些概率可选择假设的简单函数，所谓“先验”亦即由科学家考察证据之前指定给这些可选择假说的概率的“主观相信度”。

“正向概率”在两种最常见的情形中确实容易测算。这两种情形是：(a)当假说实际上蕴涵着证据时(在此情形中，假说既定的情况下证据的正向概率是一)；或者(b)当假说本身是一个统计学的或随机的假说时，其部分内容就是，所获得的特殊证据应以一确定的概率 γ 出现。在运用贝耶斯的定理时，我们遇到一个严重的困难，以致不但哲学家而且统计学家在贝耶斯公式对于理论之证实的重要性和有效性方面都发生了严重分歧。这个困难就是，缺乏一个衡量先验概率的尺度(用德·费尼提和萨维奇的术语来说)缺乏一整套“主观的相信度”。

我们暂且只讨论其正向概率能够实际地计算出来的这样一些假设。对这类假设来说，刚才谈到的方法确实是一种纯粹形式的方法。这就是说，我们能够用一个程序使一台计算机计算出在合适的“输入”既定的情况下各种假设的支持度。但是输入恐怕不仅要包括可计算的“正向概率”，也包括在一定上下文中的先验概率尺度。如果我们把该先验概率尺度看作是代表着科学家对世界的先入之见(正如“主观的概率函数”一词所意味的)，那么，看起来，这些针对方法本身的输入之一就是一整套对世界实质性的真实信仰(或相信度)。这是许多科学哲学家现今看待这个问题的方式。人们日益相信，在科学的内容和科学的方法之间无法泾渭分明地划出一条界限；科学的方法实际上如科学的内容一样，在不断地发生变化。贝耶斯定理如果确实把握了确证论的逻辑，它就通过对先验概

率函数的需要而提供了一条将科学方法对于科学内容的依赖予以形式化的途径。

我们可更抽象地说，"方法"崇拜者假定了合理性是不可分割的，但贝耶斯定理表明，事实并非如此；即使在科学领域里，在我们正在讨论的，可计算其正向概率的理论的特殊领域内，合理性也能一分为二，一是形式部分，它能够从数学上程式化，并能在计算机上编制程序；一是非形式部分，它不能被程式化，并且要依赖科学家的实际的变化着的信仰。说得婉转一些，如果合理性的形式部分足以保证获得满意的结果，那倒是不错的。如果我们能说，只要科学家进行细致的观察，收集充足的材料，然后根据贝耶斯定理，计算出支持度的话，那么，尽管这些人在开头由于主观的相信度的差异而存有分歧，但最终他们会趋于一致，于是就万事大吉。但是在这幅美妙的图景中，有两点错误需要指出。

错误之一是：即使我们能够表明，"先验概率函数"终究是要取消的，或者，即使我们能够证明，只要不断地收集更多的证据，并且运用贝耶斯定理，相信不同的概率函数的科学家最终会趋于一致，我们还是有必要要求这种会聚合乎情理地快些出现。如果相信不同先验概率函数的科学家直到被预测的现象已经出现，或者直到数百年之后才达到一致，那么，从短期来看，最终的会聚有数学上的保证这一事实就毫无用处了。长期的正当性辩护的麻烦就在于，这个长期恐怕实在太长了。凯恩斯有一句名言说："从长远来看，我们都是要死的。"第二个错误是，事实证明，先验概率函数的差异能导致分配给理论的实际支持态度的巨大差异，并且，这种差异可以相当于我们通常所理解的彻底的不合理性。

后面这层意思还可用另一种方式表达，假如一个科学家从一个"合乎情理"的先验概率函数出发，他就只会把支持度分配给那些看上去"合乎情理"的假设。如果一个人只遵从对合理性的描述的形式部分，如果他在逻辑上保持一致，只按照贝耶斯定理来确

定支持度，但他的先验概率函数却是极“不合理的”，那么他对证据支持各种假设的限度的判断就会（像科学家和平常人实际上判断这些事情时那样）是完全不合理的。形式的合理性，只承认对科学方法的形式部分，并不能保证实在的实际的合理性。

这种说法的真实性程度事实上相当令人震惊。阿瑟·伯克斯事实上已经证明，甚至存在着“反归纳的先验概率函数”，也就是说，存在着逻辑上可能的先验概率尺度。因而只要某一科学家拥有这种尺度，那么，当某一假设有了更多的证据（使用更多的证据这个术语是基于判断），那么，在一个相当长的时间内，科学家对这个假说的估计也许会越来越低。

摆脱这个困难的途径或许是：通过另外一套能够决定哪些先验是合理的（此后，我把先验概率函数简称为“先验”以符合统计学上的通常用法），哪些先验是不合情理的形式或规则来试图补充现在的对科学方法的形式解释。但是我们似乎没有什么充足的理由认定会有一组区分合理的先验与不合理的先验，比一个十全十美合理的人类总体心理的详尽描述可能更为简单的规则。看来对某种能与有关科学内容（即有关世界本质的科学）的实际人类判断和人类价值相分离的形式方法的渴求已化为泡影。而且，即使我们放宽方法的概念，以便对一个十全十美合理性的人文科学家的心理的形式化能够被称作一种“方法”，我们还是没有理由认定，一种这种意义上的“方法”应该是独立于关于美学的判断，关于伦理学的判断，或者关于随便什么东西的判断。人们以前相信科学方法不会运用于伦理学和美学方面的信念，或不会以这些信念为前提的全部理由在于相信，科学方法毕竟是一种**形式的**方法。

我对这个问题的讨论建立在这样一个假设上，即假定一种科学方法形式化的特殊方法，所谓“贝耶斯”方法，是正确的。但是，在其他每一种被尝试的方法中都会有相同问题出现。即使一个人试图将归纳方法中不像理论之确证那样“分量重的”、与培根所理解

的"归纳"更为接近的一个小小部分分离出来，也就是说，尽管一个人试图通过考察足够的例子并"预测"简单概括的真值来将一种确证该简单概括的方法分离出来，同样的问题还是会出现。古德曼已经表明，[①]对归纳预言而言，任何纯粹的归纳预测形式规则都不可能摆脱自相矛盾，甚至在一个形成规则能希望产生一致的结论之前，你就不得不事先将语言中的谓词分为"可预测的"和"不可预测的"这样两个部分。归纳的即使最基本的部分也有某一部分(即对词汇可预测部分和不可预测部分的划分)是非形式的，这一事实又一次有力地支持了我们在讨论贝耶斯定理时所提出的那个结论，即我们不能在科学家的实际信仰和科学方法之间划一明确界限。古德曼所做的在于发明了一个谓词"绿蓝"，这个谓词仅仅适用于这样的物体，它们公元2000年前被观察到且是绿色的，或者没被观察到且是蓝色的。在公元 2000 年前，每一个被考察过且被看到是绿色的物体同样也被考察且被看到是绿蓝的。任何告诉我们如下情况的预测即当我们考察了一定数目的、具有性质 P 的物体，比如说是绿宝石，我们就可以推论出："所有绿宝石是 P。"它们的形式规则都允许我们作出自相矛盾的推论："所有绿宝石都是绿色的，"并且，"所有绿宝石都是绿蓝的。"古德曼令人信服地表明了，一切在纯粹形式基础上来排除像"绿蓝"这类"古怪"谓词的企图都是行不通的。[②]

实际上，在古德曼而对培根的归纳时所碰到的困难和在贝耶斯定理方面的先验必不可少这两者之间有着密切的联系。假定在

---

① 见他的《事实，虚构和预测》，1977 年，哈克特第 2 版，第 1 版于 1954 年出版。

② 古德曼自己的解决办法是，除了考察形式，还附带考虑包含在推理中的谓词的先验预测的历史(和某些相关的事物，将如"堑壕"和"跨跃"等一起考虑)。按古德曼的意思，我们可以由此得出，一个永远预测像他著名的谓词"绿蓝"这类"古怪"谓词的文化现在也许是完全正当的——他们的推理现在也许是在归纳上有效的！

尽管我赞同古德曼的观点：与既往的实践相适合是科学的一条重要原则，但这条原则在古德曼那里还是"太简单化，太相对主义化了"。

2000年前的某个时刻，科学家们不得不从中择一的两个假设是："所有绿宝石都是绿色的"以及"所有绿宝石都是绿蓝的"。假定，相关的证据是，大量绿宝石被考察且都被发现是绿的(因而也都被发现是绿蓝的)。如果科学家用贝耶斯定理来计算两个假设的支持度，那么结果将会证明，他或者能为那个正常的假设(所有绿宝石都是绿的)找到极高的支持度，或者能为那个反常的假设(所有绿宝石都是绿蓝的)找到一个极高的支持度，或者为这两种假设找到同样的支持度，具体情况取决于科学家的先验。如果一个人的主观概率尺度将这样一个较高的先验概率确定给"所有绿宝石都是绿的"，而不是给"所有绿宝石都是绿蓝的"，那么事实上，他的行为就像是他正在预测"绿的"，而不是"绿蓝的"。从贝耶斯派的观点看来，在一个人能够作出一个归纳之前，要对哪个性质谓词是可预测的，哪个性质谓词是不可预测的作出一个决定，只需要一个先验的特例。

波普尔已经指出，在可选择的假设中，一个人应该选择最可错的假设。但是结果证明，他对可错性的形式衡量将依据人们选择某语言的哪些谓词作为初始谓词而导致不同的结果。不管一个人是像波普尔那样，认为科学家就是试图发现还未被排除的最可错的假设，还是以更为常规的方式认为，科学家试图计算假设的支持度，对一个对应于古德曼关于某些谓词是可预测的，其他谓词是不可预测的非形式要素的需求，或者对应于贝耶斯派的先验的认可的非形式要素的需求依然是必要的。

关于这一点，读者或许会奇怪，如果没有科学方法这种东西，或者，如果这种方法即使能够形式化也依赖不能形式化的输入，那么我们如何解释科学的成功？不可否认，科学一向是一个令人瞩目地成功的事业。我们很容易感受到，这种成功的原因必定与某种事实有关，这一事实就是在科学家着手获得知识的方式和普通人在前科学时代着手获得知识的传统方式之间的区别。这完全错了吗？

答案是否定的。我们必须作出的选择并不是:要么认为科学由于遵从某种精确的形式规则系统而获得成功,要么科学之成功纯属运气。从15世纪开始,科学家和哲学家着手提出一套新的方法论准则,这个过程至17世纪达到鼎盛。这些准则并非精确的形式规则。这些规则的确要求运用非形式的合理性,如理智和常识等;但是科学研究过去是现在仍旧是由它们说明的。简言之,存在一种科学方法,但是这种方法预设了一些先验的、合理性概念[①]。再说一遍,充任无所不包之物的、充任合理性定义的并不是一种方法。

17世纪最重要的方法论大师之一是物理学家波义耳。早在17世纪之前,物理学家在实际地从事实验和仅仅描述确证他们或多或少是在先验基础上相信的理论的思想实验之间没有作明确的区分。不仅如此,科学家也没有意识到发表不成功的实验报告的必要性。一句话,实验在很大程度上被看作是基于演绎和先验根据而相信的学说的图解而不是作为支持或反对理论的证据,波义耳写过关于实验程序的小册子,他强调必须在思想实验和实际实验之间作明确区分,并且强调要详尽地描述一个人所作的全部实验尤其包括不成功的实验。波义耳本人是哲学家培根的门徒,他无疑被培根的归纳观点所影响,从而对这些规则的重要性予以高度肯定。然而事实上,在形成物理研究程序方面,波义耳提供的具体要领较之培根提出的对归纳程序的较为抽象、较为公式化的辩护,其作用更为重要,或至少同等重要。

从试图先验地建立一个理论转向试图通过从理论中引出可检验的结论并从事实验来检验理论,这无疑是一个方法论上的转向。然而,正如我们已经看到的,我们决不能简单地将是合理的与仅仅因为理论受到精心从事的实验的支持就去相信理论等同起来。一

---

① 穆勒本人以一种极勉强的语调承认这一点。他写道:"只要我们允许绝对的白痴插手,"我们就不要指望归纳方法发挥效用(《功用主义》,第2章)。

方面，即使在科学中，也不是永远能够从事可控制的实验。有时，你只能依靠被动的观察，而不是依靠“实验”一词所蕴涵的那种主动的干预。而且，如前所述，甚至当你为了从各种理论中进行选择而做了一个精心的实验，实验结果支持各种可选择理论的程度的估量，依然是一件完全非形式的事情。

和我们一贯的看法相反，波普尔反复争辩说，有一种独特的科学方法，它可以陈述出来，而我们则应该仅仅依靠它来发现世界的本质。

但是波普尔确实以为，存在着一些比科学合理性更为宽泛的合理性概念，而且它确实适用于作出伦理上的决定。

在其颇具影响的著作《科学发现的逻辑》(《研究的逻辑》)和后来的一些著作中，波普尔认为，科学的方法在于提出各种“高度可证伪的”理论，含有冒险猜测的理论。然后，我们开始检验这些理论，直到仅仅一种理论幸存下来。然后，我们把这幸存的理论当作能暂时应付一下的东西而接受，然后是整个过程重复出现。既然除此幸存的之外，其他所有理论都是根据演绎而加以排除的(当一个理论蕴涵着一个肯定已被证伪的猜测时，这个理论就该排除)，波普尔宣称，贝耶斯定理便毫无用处，当然，支持度的估量也不会出现。

波普尔的观点有这样一个问题：它不可能检验所有可证伪性很强的理论。例如像这样的一种理论：如果我在头上罩一面袋，轻轻敲 99 次桌子，一个魔鬼将出现，它的可证伪性当然是很强的。但我肯定不会劳神去检验它。就算我愿意去检验它，我可能要想到 10^{100} 个同样的理论，这样，我一生的光阴，甚至人类的漫长岁月都不足以将它们检验完。因此，出于逻辑上的原因，有必要在方法论的基础上选择数量极少的我们愿意实际地去检验的理论。这意味着即使在波普尔的方法中，也有一个类似先验选择的东西存在。正如我在前面已评述过的，就连波普尔对可证伪度的计算也同一个

人以自己的语言中哪些谓词作为初始谓词息息相关，在这个意义上，甚至连可证伪性概念也需要一个先验的决断，这个决断和古德曼关于某些谓词是"可预测的"、某些是不可预测的这一决断有类似之处。我想把这些与我们当前讨论关系不大的技术性问题暂搁一边。即使波普尔的方法是不完备的，需要用一个我们眼下还无力加以形式化的、更为直观的方法去补充，我们不也可以说，他的理论提出了对科学合理性的一个必要条件吗（如果谈不上充分条件的话）？简言之，我们难道不可以说，对科学理论的可接受性来说，其必要条件是这个理论能够经受得住波普尔的检验吗？波普尔的检验本身也许包含着对待检验理论的先验的选择，这个先验的选择本身是非形式的，我们对之没有一个算法可用。但是我们仍旧可以坚持说，除非起初选择的一套理论都是直观上高度可证伪的，除非所有我们没有接受的理论都是后来经过精心从事的实验而加以拒绝的，否则我们不能接受任何理论，简言之，我们难道不能说，我们不能接受任何理论？简言之，我们难道不能说，我们应该给予科学家的忠告是：按波普尔的忠告行事，在波普尔的方法不能形式化的地方，对波普尔的准则该作何解释，依靠你的直觉？还有，波普尔的方法尽管在某种程度上既含糊又非形式化，但可能不仅穷尽了科学合理性概念的外延，而且穷尽了所有认知的合理性概念（也就是说，一个陈述是有据地可断定的，或者，是合理地可接受的，当且仅当它蕴涵在一个基于波普尔的检验方可接受的理论之中）的外延。事实难道不可能是这样的吗？回答是，这样的合理性观念哪怕对科学而言也是太狭窄了。一方面它也许排除了对达尔文的自然选择的进化论这个所有科学理论中最成功、最受推崇的理论的接受。对这个后果波普尔是能够心安理得地接受的，但科学共同体则显然不能。自然选择理论不是高度可证伪的，它并不蕴涵像如果它们出错，则该理论就被拒绝这类明确的猜测。我们之接受自然选择理论，并非因为它经受了波普尔的检验，而是因为它对大量的材料

提供了似乎有理的说明；是因为它在启发人们提出新的理论，在把遗传学、分子生物学等的各种进展联成一体方面一直是卓有成效的；还因为人们实际已经提出了其他理论，从背景知识的角度讲，要么已被证伪，要么几乎是完全不可能的。简言之，我们把达尔文的自然选择的进化论当作皮尔士所说的不明推理来接受，或者当作眼下所说的导致"最佳说明的推理"。这恰恰是波普尔希望从科学中驱逐出去的那种推理。但是科学家并不打算因此接受波普尔的忠告，在一些理论对不计其数的材料提供了极好的说明的情况下，在该领域中没有其他言之成理的说明可资利用的情况下，去放弃这些没有很强可证伪性的理论。确实，正如我在另一篇文章中[1]指出的，波普尔哪怕对于经典物理学的理论本身所具有的很强的可证伪性的程度，也作了夸大。

因此，通过承认"导致最佳说明的推理"是科学推理的合法形式（即使推出的"最佳说明"不具备波普尔所说的很强的可证伪性），我们进一步淡化对于科学方法的描述，科学方法现在成了一个意义极其含糊的东西。[2] 但是无论如何，从我们前述归纳逻辑中的形式的后果来看，这是我们原来就意料到的。我们是否可以说，这个描述得如此含糊的"科学方法"现在已经包括了所有科学方法呢？我们是否可以说，即使根据这样含糊的一个描述，任何价值判断都不能借这个方法得到证实或确证，乃是一个百分之百的事实呢？回答是，如果科学方法仅仅被说成是"尽可能仔细地做实验和观察，然后作导致最佳说明的推理，排除那些被判决性实验所证伪的理论"，那么你就不能理解什么是不能被一个描述得如此含糊的科学方法所证实的。例如，假定我想证实"约翰是个坏人"这一陈

---

① 参见"理论的确证"，收于拙著《数学，物质和方法》。

② 要不，我们可以把"科学方法"一词局限于专指实验过程准则的有意识的运用，如我在《意义和道德科学》中所提议的那样，而不再企图使科学方法具有能包容我们称之为知识的一切事物的伸缩性。

述，我也许作如下论证："以下是已观察到的事实，约翰是粗心的人，是极自私的人，又是极冷酷的人；凡粗心、自私和冷酷的人显然是一个坏人；所以约翰是个坏人。"这一推论在两个方面会遭到一个持有"价值判断"不能"科学地加以证实"这种观点的人的反对，反对者会针对这一推论的最后一个步骤：即，从约翰是冷酷、粗心和自私的人推出约翰是个坏人。无可否认，这是一个概念上的步骤。这就是说，在是冷酷的，是粗心的，是自私的和是道德上坏的之间有一个概念上的联系。[①] 当然，假如在道德谓词之间根本没有概念的联系，那么这就是一个无效的推论。但是一个人为什么要相信，在道德谓词之间没有概念的联系呢？也许有人会说，在一个推论中，运用所谓"概念的"步骤本身是不科学的。但是，要说在科学自身中根本没有这一步骤，这种观点肯定是站不住脚的。例如，如果我从牛顿对太阳系的描述中推论出"引起潮汐的是月亮引力的吸引"这一陈述，那么我就正在运用我关于在力的陈述和具有A引起B这种形式的陈述之间有着概念的联系这种非形式化知识。"引起"这个词在牛顿对太阳系和潮汐的描述中甚至都没有出现；但仅仅根据对牛顿理论的理解，我就知道A施加于B的引力是由A（的质量）引起的。

当然，如果我们把科学的方法描述为从自身是价值中立的语言的"观察陈述"出发作出导致最佳说明的推理，或诸如此类的东

---

① 真理或推论是一种我们称之为"概念上的"事物这一事实并不意味着它的性质一定是纯语言学的（亦即是由于任意的语言约定而真的）。许多具有不同思想倾向的哲学家已看到，概念、事实和观察是相互依赖的。正如我们在第六章里评述过的，概念是由我们观察到或直观到的东西形成的。反过来，概念又形成我们能观察到和直观到的东西。就这点而论，本文中包含有"好"字的推论就与包含"有意识的"这字的以下推论非常相像。"约翰在有理智地说话，在得体地做事，并且对所发生的一切有反应，所以，约翰是有意识的。"这里，"有理智地说话"、"得体地做事"、"对所发生的一切有反应"等就是将约翰归结为"有意识"的明显原因，恰如粗心、自私和冷酷等，是将约翰归结为道德上坏的明显原因一样，这就是概念上的联系。

西，那么我们就可以把“约翰是粗心的”和“约翰是自私的”这类陈述排除在“观察陈述”之外(尽管在特殊情形下，对这些陈述要比对——比方说——关于一个对象是不是紫红色的陈述更容易得到一致的意见)。但是，这类陈述频繁出现于各种著作，例如，历史学家的著作中。历史、临床心理学和日常语言描述可以完全避免“体贴”、“自私”这类词语是不大可信的(而且，从何处划这条界限也是个大问题：“固执”是价值中性的吗？“发怒”是价值中性的吗？)。然而无论如何，将合理性等同于这样规定的科学合理性，也许是用本身需要论证的价值判断的认知性质来进行论证，这等于说，这些判断之所以是不能合理地确证的，是由于它们是价值判断，因为合理性已被定义为纯粹是由原始中性的观察组成，并且从中性价值的前提推出结论。但是我们为什么要接受这样一个定义呢？

第九章　价值，事实和认知

在第六章，我已论证过："每一事实都含有价值，而我们的每一价值又都含有某些事实。"其论据概要说来是，事实（或真理）和合理性是相互依赖的概念。一事实也就是可合理地相信的东西，或更确切地说，事实（或真陈述）的概念，是对于能合理相信的陈述的概念加以理想化的结果。"可合理地接受的"和"真的"是两个相互在对方身上显露其真相的概念。我还论证了，是合理的不仅意味着具备合理的可接受性的标准，而且意味着具备相关性的标准，并且论证了，我们所有的价值都包含在我们的相关性标准之中。决定说一世界图景是真实的（或是从我们现在的眼光来看是真实的，或是"像其他任何事物一样真实"的），并且说它回答了某些相关的问题（而不仅仅我们能够回答这些问题），这依赖于、并且揭示了我们整个的价值承诺体系。一个没有价值的存在也就无所谓事实。

通过考察最简单的陈述，我们至少可以间接地看到相关性标准是怎样包含价值的。以"猫在草垫上"这句话为例。如果某人在一个特殊的语境中实际地作出了这个判断，那么他就运用了概念元素"猫"、"在……之上"和"草垫"。这些元素都是由一个特定的文化所提供的，其出现和普遍存在揭示了有关该文化，以及差不多每一文化的价值和兴趣的某些方面。我们之所以有"猫"这一范畴，是因为我们认定世界之划分为动物和非动物是有意义的，还因为我们对一个已知动物的种属也有浓厚兴趣。有一只猫而不仅仅是一

个物在草垫上，是和这种意义和兴趣相关的。我们之所以有“草垫”的范畴，是因为我们认为非动物之划分为人造物和非人造物是有意义的，因为我们对一个特定的人造物的用途和本质有进一步的兴趣。因此，猫是在草垫上而不是在某物上，这是和上述意义和兴趣相关的。我们之所以有“在……之上”的范畴，是由于我们还对空间关系发生兴趣。

请注意我们得到的结论：我们列举了“猫在草垫上”这个可想象的最简单陈述，从中发现，这个陈述要在一定语境中具有相关性的前提是动物/非动物，用途和空间等范畴是有意义的。对那些无意将这些东西认定为恰当的范畴的人来说，“猫在草垫上”这个说法，也许会和少男少女在热恋时的窃窃私语中说出：“这间屋子里的六边形物体的数目是76”一样不合理。

不仅仅有关我们价值体系的极普遍事实显示在我们的范畴（人造物，种属名称和空间关系所用词汇）中，而且，正如我们在第六章中已看到的，许多更具体的价值（如感受性和同情心）也在我们使用一些特定分类词（如“体贴”、“自私”）的过程中显示出来。重申一遍，我们的相关性标准依赖于并且揭示了我们的整个价值系统。

这种对相关性的讨论，同上章提出的那个问题（“何为合理性的价值”）有直接的相关性。如果合理性是这样一种能力（或确切些说，一种诸能力的整合系统），它使拥有这种能力的人能够确定什么是能相关地提出的问题以及什么是有根据地接受的答案，那么它的价值就显露无遗了。但是这样一个合理性概念就像相关性概念本身一样，也是含有价值的，这一点无须争辩。

或许有人反驳说，你这是把原本独立的因素硬扯在一起，因此其中必玩有花招。我只谈到合理的可接受性和相关性这两个因素，这一事实本身也可以被说成是证实了类似于事实/价值二分法的这种东西的长期性和恒久性。根据这个持反对意见的人的看法，一个合理的人，也许是一个能辨别什么是有把握断定的、什么是没有

把握断定的人。他们会说，一个人把什么选作有趣的，或重要的，或相关的东西，也许同评价他的性格，甚至他的心理健康状况有关，而与其认知的合理性无关。

然而，可接受性和相关性在任何实际语境中都是相互依赖的。使用任何词——不管这个词是“好”、“有意识”、“红”，还是“有磁力”——都使人卷入一个历史过程，一种观察、概括、实践和理论的传统之中。

这也使人卷人解释该传统的活动之中，卷人使传统适应新情境、被延续和被批评的活动之中。对传统可作各种解释，但是如果一个人将自己完全置于某个词所属的传统之外，则他将根本无法运用这个词。置身于传统之内无疑会影响一个人对“合理的可接受性”的看法。如果毕竟有一种能用来证实任何陈述的方法（无论它包含什么概念）的话，那么有人提出的把证实陈述的能力与对于一套相关的概念的掌握区分开来，也许是站得住脚的。但是我们已经看到，我们没有理由接受一种单一方法的神话。

两 要 素 论

眼下我们对合理性的种种直观似乎是互相冲突的：可以断定没有一种哲学理论看上去能把这些直观都协调起来。一方面，如果我们说，我们从未把目的判别为合理的和不合理的，当然是不正确的；另一方面，当我们面临“合理的纳粹”这样一种假设情形时，尽管我们知道纳粹是在道义上令入反感的，并且把这种经过周密策划而又深思熟虑的目的体系说成是不合理的，但是我们却无法证明这种批评的正当性。

据说有可能解决此类问题的一种途径是这样的：由于假定了事实/价值的二分法，一旦我们碰到一个只热衷于搞清人的头上有多少根头发的人，我们便完全有理由根据他对事实的这种不恰当

的(当然,所谓恰当的意思是什么也是一个问题)理解,而指责他是不合理的。纳粹只是在价值上和我们有分歧,这就是他何以不是不合理的缘故。介于两者的情况,也许可以按照威廉斯提供的线索加以讨论。尤其是,人们可以提出一种关于日常语言的道德-描述词汇同时具有两种"意义要素"的理论来对付我们反对方法崇拜的那个论证,即除非我们用循环论证,否则我们就不能从"观察证据"中排除诸如"约翰会体贴人"这类描述判断。这两要素之一是事实要素。对体贴,无疑存在着公认的标准,而且"约翰会体贴人"这句话传达的是约翰符合这些标准[1]的信息。然而还有一个情感意义要素。"约翰会体贴人"传达了对约翰行为的某一方面的"前态度"。按这种观点,能成为合理的,是对"约翰会体贴人"这一陈述的事实要素的认可;而认可那种情感意义要素,即具有那种"前态度",既不是合理的,也不是不合理的。

此处的"事实的"指什么,取决于哲学家个人对世界的构成物所持的看法。对一个唯物主义哲学家来说,任何陈述的意义中的"事实"要素必须存在于可用物理科学词汇表达的陈述之中。但是随即会出现一个使现象主义者苦恼的困难。

我们记得,现象主义是这样一种学说,即,一切有意义的陈述都能够毫无保留地译成关于感觉的陈述。而"物理主义"(我们正在讨论的那种唯物主义常常被人这样称呼)是这样一种学说,即,任何陈述的全部"事实"意义能够毫无保留地翻译成物理学的语言。

---

① 但此处"标准"的含义也成问题。假若"约翰会体贴人"这一论断从描述上说是真的(即它的"事实要素"是真的)当且仅当多数说话者会同意"约翰会体贴人",那么,从这样一种对事实要素的分析,或许会推出,不可能有这种情况,即尽管多数说话者反对,一个人仍旧被某一敏感的法官正确地归入"体贴"一类的人。对"事实要素"的这种解释也许只相当于宣称,所有真理(至少关于标准的真理)必须是"公共的"。但是除非一个人是一个多数主义者,否则人们为什么要相信这一点呢?(正如我们在第五章看到的,所有真理是公共的这一论断是自相矛盾的。)

这种学说看来又是不能成立的。

我们不考虑道德描述陈述，而考虑一个心理学陈述，比方说，“X 正想着维也纳，”来看看为什么物理主义不能成立。很显然，对一个正在想着维也纳的任意的有机体来说，即使存在着可表述为大脑状态之类的必要条件和充分条件，要说这些条件是什么，也将求助于一个完善得难以想象的神经病学(或机能心理学)理论。这个陈述的真值条件是依赖于语境的，相关于意趣的，并且是含糊不清的。没有理由认为(哪怕是在原则上)存在着用物理主义语言表达的这样一种确定的表达式，这个表达式之适用于 X，当且仅当 X 正想象着维也纳。不仅不可能有一个与日常语言陈述“X 正想着维也纳”确切对应的用物理主义语言表述的陈述，而且就算这样一个对应陈述真的存在，其对应也是基于一个经验理论或一组本身尚未被人知道的理论(或许它们太复杂以致人类决无可能知道它们)，而且这些理论肯定不是“X 正在想着维也纳”的意义的一部分。简言之，若是以为“X 正想着维也纳”意指“X 正处于如此这般一个(物理学上或功能上已说明的)大脑状态”，是错误的。

对“X 正想着维也纳”是适用的，对任何其运用条件与支配物理概念的这些条件不能默契配合的日常语言的谓词也是适用的。“X 会体贴人”—— 乃至“X 是棕色的”，“X 暴跳如雷”，“X 是一个人”——同样不能翻译成物理主义的语言。这意思是说，如果“X 会体贴人”确有两种意义要素，那么我们能给予这陈述的“事实意义”的唯一表述是，这句话是真的，当且仅当 X 会体贴人。这样一个事实要素的概念未免失之琐碎。

说两要素论之无效，并不否认“X 会体贴人”通常有某种感情色彩。只是它并非永远具有这种色彩。正如我们在第六章已指出的，我们可以将“X 会体贴人”用于各种目的，用于评价、描述、说明、预言等等。但是对一个陈述可具有的各种不同用法加以区分，并不要求我们否认“X 会体贴人”这样一个陈述的存在。

摩尔和“自然主义的谬误”

韦伯关于“价值判断”不能加以合理确证的论断是当前事实/价值二分法的根源。然而这一二分法却被G·E·摩尔强化了(这和他本人意愿相左)。当罗素、凯恩斯以及布隆姆伯里小组后来的其他一些成员还是稚嫩的大学生时,摩尔就在论文中论证道,善是一种非自然的性质,即,是一种完全处于自然科学的物理主义本体论之外的性质。他对自然主义的辩护结果适得其反。摩尔的弟子们可能相信了他的话(尽管至少罗素对他即将表示怀疑),认为存在着非自然性质之类事情,但是后来的自然主义型的哲学家却倾向于认为,摩尔在此提出的是对于有价值性质之类东西存在这种观点的旧谬反驳。在30年代,史蒂文森和逻辑实证主义者们开始提出“情感论伦理学”。这种理论是说“X是善的”,意味着“我赞赏X,并予以仿效!”或诸如此类的说法。价值性质开始从认识论的基础上,然而更多地是从本体论的基础上加以拒绝。正如麦基[1]最近所说的,应当有价值态度这种东西,这与自然科学是和谐共存的,而如果居然有价值性质这种东西,则与自然科学不相容。麦基指出,价值性质是“本体上可疑的”,即,它们在其存在上有近于荒唐的神秘性质,被科学所开化的人不会继续相信它们的存在了。

摩尔对善不可能是一个物理主义性质(一个“自然”性质)的论证是,如果善和“有助于获得最大限度的总体效用”有同样性质(或者和其他什么自然属性——物理的,功能的,只要你愿意去代换——有同样性质),那么

(1)“这个行为不是善的,尽管它有助于获得最大限度的

① 《伦理学,创造正当与非正当》,企鹅出版社,1977年版。

总体效用”

就是一个自相矛盾的陈述(而不仅仅是假的陈述)。

殊不知,即便一个功利主义者也不会宣称(1)式是自相矛盾的。摩尔指出,这表明,尽管善和有助于获得最大限度的总体效用是两个相关联的性质,但它们不可能是相同的性质。

然而,摩尔的论证依赖于一些我和其他语言哲学家今天会拒绝的假定。首先,他含蓄地否认了,可能存在着一种性质的综合同一性。但正如我在第四章指出过的,这会排除掉像温度的量和平均分子动能的量是同一个量这类公认的科学发现。(有人会利用摩尔的“证明”来说明,事实上温度必须是一个“非自然性质”。因为当一个人说“x 有温度 T 但 x 并没有平均分子动能 E”——在此,E 是对应于温度 T 的动能的值——时,他并不自相矛盾。尽管这个陈述从经验事实上讲是永远假的。所以摩尔会得出结论说,温度仅仅是与平均分子动能相关联的;这两种性质不能严格等同。)事实上,摩尔把性质和概念合二而一了。有一种关于性质的看法,根据这种看法,两个概念(比方说温度和平均分子动能)是有差异的这一事实,根本不能解决相应的性质是否也有差异的问题。(发现有多少基本的物理性质存在不是在概念方面有所发现,而是在世界方面有所发现。)“善”这个概念与其他任何物理概念可以不是同义的(在古德曼看来,道德-描述语言和物理语言毕竟是极不相同的“文体”),但是,这并不能得出,是善的和是P——如果P是某种恰当的物理主义的(或更确切地说,功能主义的)P 的话——不是同样的性质。一般说来,对一个性质的用直接方式习得的词汇(如“有高温”)与该性质的理论定义不是同义的。把握温度为何物(某个哲学家会说,何为善),需要一个经验和理论的研究,而不是语言分析,不是对意义的反思。

在我引进了“性质的综合同一性”概念数年之后,克里普克扩

充并阐明了我正在讨论的观点。他提出只能经验地获得的“形而上学的必然真理”的观念,“认知上偶然的必然真理”① 的观点,这种观点在语言哲学中广为流传。联系温度/动能的例子,克里普克的看法是,如果某人描述了一个逻辑上可能的世界,在这个世界上,人们有冷热感觉,有感到热的对象,并且,热和冷的感觉都由与平均分子动能不同的机制来解释,那么我们不说他描述了一个在其中温度不是平均分子动能的可能世界,而是说,他描述了一个在其中某些不同于温度的机制使得某些对象感到冷和热的世界。一旦我们承认(在真实世界上)温度是平均分子动能这个综合同一性陈述,就没有什么可以被称为其中温度不是平均分子动能的可能世界。

在每一可能世界中为真的陈述历来被称为“必然的”陈述,而某物在每一可能世界中都具有的性质历来被称为“本质的”性质。克里普克用这种传统术语要说的是,“温度是平均分子动能”乃必然的真理,尽管我们不可能先验地知道这一点。这个陈述是经验的,然而是必然的。或者用另一种不同的术语来表述这同一内容,平均分子动能是温度的一个本质性质。我们是通过经验探究才发现了温度的本质的。克里普克的这些思想对语言哲学、形而上学和数学哲学有广泛的影响,并运用到他们正在攻击的摩尔的论证之上。摩尔从(1)式只能是偶然地虚假的这一事实推出P(对某些合适的自然性质P来说)不可能是善的一个本质性质,这正好是新的必然性理论所拒绝的。一个人能从(1)式不是自相矛盾的这一事实中有效地推论出来的,充其量只是“善”与“有助于获得最大限度的功利”不是同义的(对物理主义世界中的任何词汇P来说,是不同义于P)。从这种词的非同义性中,得不出任何有关性质的非同

① 《命名和必然性》,哈佛,1980年版(此讲演最初于1970年在普林斯顿大学所作)。

一性的结论，也得不出任何有关善的本质的结论。

露丝·安娜·普特南指出，关于善不能是一种自然性质的另一种常见的论据也不成立。[①] 这个论据是："X 是善的"含有"感情色彩"，"表达一种前态度"，等等。

露丝·普特南指出，这种推理的错误之处在于，许多描述性谓词天生拥有一种感情色彩，在我们的文化中，"把饭撒了一身"（这人是个漏下巴），有一种很强的否定的感情色彩，尽管这个短语从字面上讲是一个性状描述。任何代表人们在某一文化中认为有价值（或无价值）的某种东西的词，都会拥有感情色彩。

"善"[②]这个词，就其道德含义上讲，可适用于众多事物。其中有些——例如，好的心境——可以被自然地予以好评：它可以是人们予以好评的那种心境本身的一部分内容。我不想反过来推论说，凡在这个意义上被自然地予以好评的心境就是善的。这明显是不对的。假定我们对善作这样一个限定，以至于只有被自然地予以好评以及没有什么正当理由予以恶评的东西（正如有理由对某些由药物引起的自然地予以好评的状态加以恶评一样）才可称为是"善的"。

这样的话，人们会期待某物是善的这一陈述，由于善这一性质的本质而具有正面的"感情色彩"。即使不是效果论者（即认为有足够好的效果的任何事物便是善的），人们称一个行为是善的最通常的理由也毫无疑问就是它能引起的效果。其中之一也许就是，它最终促进了能自然地予以好评的状态或境况，这里，正是性质的本质再次解释为何性状描述最终具有"前"感情色彩。

麦基为他的善是本体上"可疑的"这一观点辩护的前提是这样

---

① "维特根斯坦论伦理学演讲评论"，载《第4届维也纳国际维特根斯坦讨论会、语言、逻辑和哲学论文集》，1980年。

② "good"一词既可译作"善"，也可译作"好"。在日常用语中常译作"好"。——译者

一个假定：若无对某物的"前"态度，则不可能知道某物是善的。这无异于用假定情感主义来证明情感主义。地狱中的魔鬼通常被描绘为用反面的感情色彩来使用"善"这个词（一个魔鬼兴许会说："此人有一种趋于道德之善的可悲倾向"）。和麦基相反，我认为这种用法并不是语言学上不恰当的，或者包含了什么矛盾的。难道我们没有听到过人们说："我知道做那事是不好的，但我不在乎？"正如福特指出的，一个人可以说"我并不想做一个正人君子"，从而置道德于不顾。这些说法所要表明的可能性是：虽然在语言的描述用法和规范用法或褒义用法之间确有一个区别，但这个区别不是词汇的一个简单函数。"描述"词能用来赞许或责备（"他是个漏下巴"），而评价词也能用来描述和说明。（考虑下列对话："约翰能这样做，看来一定是个大好人。""不，金无足赤，他也不是一个道德完人。不过他有一种出人意料的自我牺牲精神倒是真的。"这里的道德语言被用作一种说明性功能。）重申露丝·普特南的观点，从"是善的"被用作褒奖这一事实，不会得出任何关于善不是一个性质的结论。

露丝·普特南教授指出，麦基的推论在某些方面还是有可取之处的。某些道德表达式不可否认确实具有一种内在的行为倾向。"应该"、"应当"、"正当"、"必须"等就是这些"引导行为"的词的最基本例子。她指出，"是/应当"的问题，不同于事实/价值的问题。"我不想做我应该做的"听上去比"我不想做一个正人君子"要古怪得多（而"我不想做我必须做的"这话听上去神经有毛病）。

麦基指出，没有什么物理性质与行为（或与赞成一个行为）有内在关联，并且得出结论说："是一件应当做的事"等等，是"从本体上可疑的"。但是除了依赖于对世界的物理主义观点是唯一真实的理论的假定之外，这个论证的根据并不充足。因为某些认知性谓词（例如，"可合理地接受的"，"有正当理由的信仰"）也是引导行为的（广义地使用"行为"一词，从而接受一个陈述也算是一个行为）。一个人可以说，"做 X 是一件好事"、"Y 有大量证据"，却不保证去做

(或建议做)X或接受Y;但是若一个人说,"在这种情况下做X是一个正当的行为"或"相信Y是完全有正当理由的",那么他就准备去做(或建议做)X并且去接受Y。"有正当理由的"(在信仰的情形中)和道德含义的"正当"一样,具有引导行为的特征。

如果我们现在步麦基之后尘,断言不存在有正当理由的这样的属性,只存在"寻找正当理由的态度",那么我们便向相对主义靠拢了。在这种可怕的后果到来之前,我们应暂且驻足,看看为什么引导行为的谓词对一个坚定的物理主义者似乎是"本体上可疑的"。

在第二章,我已论证过,对一个坚定的物理主义者来说,指称本身似乎就应该是"本体上可疑的"。如果真正存在的唯有物理主义意义上的性质和关系,那么指称要能存在的话,一定是一个物理主义意义上的关系。然而正如我们所看到的,问题在于"候补者"的过剩。可接纳的指称关系的数目是无限的(而且,所有这些关系都将是物理主义的,如果我们把集合论当作世界的自然主义理论的一支,至少也是自然主义的)。假定其中之一是唯一的指称关系,那么这一事实本身也许是某种极为奇怪的、终极的形而上学事实。

使这一事实如此奇怪的原因原来是我们将某种中立性和无意性加之于我们的自然界的概念之中。自然据说是没有兴趣,没有意向,没有观点的。如果这个说法是正确的,我们何以可能形而上学地挑选出一个可接纳的指称关系呢?

还是这个自然界的无意性,使得"是正当的"和"有正当理由的信仰"这类引导行为的谓词变得"可疑"起来。如果一个物理主义性质P等同于道德上的正当或认识论的有正当理由,那么这些谓词也会可疑的;就像如果指称是一个物理主义的关系,这些谓词也将是可疑的一样——这两者的可疑性出自同一原因。这些假设好像在说,自然界本身在道德领域里具有价值,在意义领域里具有指称意向。

因此，我认为，就摩尔坚持"善的"、"正当的"（还包括"有正当理由的信仰"、"指称"和"真的"）不等同于物理主义的性质和关系而言，他是对的（尽管他的推论是不足取的）。这不是说，善、正当、认知上有正当理由、指称和真理并不存在，而是说，一元的自然主义（或物理主义）是一种不恰当的哲学。

再论"合理的纳粹"

在前文中，我们颇感棘手的是，我们不知从何入手论证假定的"完全合理的纳粹"有一个不合理的目的。问题大体是这样的：我们太轻率地将（作为有着一种或几种世界观的人的）纳粹的合理性问题与纳粹的目的的合理性问题等同起来。如果纳粹"之中"不存在我们可以诉诸的目的，那么说"卡尔有一不合理目标"，从而判断局势，确实显得很荒唐。尽管这是我们的结论之一，但毫无疑问，我们首先要说的则是，卡尔有一个罪恶的目标，而不是一个不合理的目标。

但是，若真要讨论卡尔的合理性，需关注的问题是其信仰和论证的不合理性，而不是其目标的不合理性。

首先，假定卡尔宣称，纳粹的目标是道德上正当的并且是善的（正如纳粹分子在事实上而不只是在哲学家所举的例子中通常宣称的那样）。那么，事实上他是在信口胡诌。他将会肯定各种虚假的"事实"命题，例如，民主制被"犹太阴谋家"操纵。此外，他还会提出缺乏有效论据的道德命题，例如，谁是雅利安人，谁就有义务迫使非雅利安人服从"统治种族"。我所诉诸的"有效论据"概念是通常道德讨论所固有的。但这个概念是适当的，如果纳粹设法在通常道德讨论的范围内来为其正当性辩护的话。

另一方面，假定纳粹像我们假定的超边沁主义者那样断然拒绝通常的道德概念。我曾论证道，一个拒绝了通常道德概念的文

化，或代之以出自别的意识形态和道德观念的概念的文化，以我们现在的眼光来看，也许会丧失准确而深刻地描述通常人际关系、社会事件和政治事件的能力。当然，如果这个不同的意识形态和道德观念比我们现有的道德体系要优越，那么这种替代也许是好的、明智的；但是，假定它们是有害的意识形态和道德观念，特别是，假定它们是有偏见的、邪恶的，那么结果将只能是对人际关系和社会事实的不恰当的、肤浅的、令人反感的表述。诚然，“不确切的、肤浅的和令人反感的”这些词反映着一些价值判断。可是我早已论证过，概念系统的选择必然反映价值判断。而且，认知的合理性尽在概念系统的选择之中。

即使个别的纳粹尚未丧失使用现有道德描述词汇的能力，即使他头脑中某个地方还保存着一些旧的概念（正如某些学者依旧熟悉并能使用中世纪“骑士”的概念一样），这些现有的道德描述概念，如“体贴”、“同情”、“正义”、“公平”等等，依然不会是他用来支配其人生的概念；它们不会真正出现在他的世界观之中。

我想再次强调的是，我并不是说，充当一名纳粹，其邪恶之处在于导致一个人具有一个有偏见的、不合理的信仰。充当纳粹的邪恶之处在于导致他所做的一切。纳粹是罪恶的，并且他也有一个不合理的世界观。有关纳粹的这两个事实是有联系的、不可分割的。但这并不是说，纳粹的罪恶首先因为他有一种不合理的世界观，而是这世界观的不合理性构成了罪恶。不过在我看来，这里仍旧可以区分目标的合理与不合理：当一目标是这样的，即，一个人一旦接受并且追求这个目标，他要么会为这个目标提供一个荒唐而虚假的论证（如果他在我们正常的概念系统范围内从事正当性辩护工作的话），要么就只能为表述通常的道德-描述事实（例如某人有同情心）而采纳一个不合理的系统作为替代物，这时，我们就有权称这个目标为“不合理的目标”。在描述和考察道德相关事实时运用一个合理的概念系统与拥有某些种类的一般目标而不是其他目标

之间，毕竟是有关联的。

“但是，倘若纳粹对成为一个纳粹，除了‘这本身就是我喜欢做的’之外，再无其他理由，那又怎么样呢？”这个问题问得合情合理，不过此处合情合理的回答无疑也是正确的回答：在此情形中，纳粹的行为，除了是罪恶的之外，还是完全任意的。注意，“任意的”属于我一直在使用的道德-描述词之一。这类词可不改其指称而用来评价（在评价纳粹的例子中是责备）、描述（“约翰相当任意地决定了改换工作”）、说明（或者指出不能给出某种说明），等等。不错，当我适才说，（在我所描述的例子中）卡尔也许是很任意地决定成为一个纳粹，我原本是在描述而不是在评价。若细究起来，我所做的许多事情都是任意的。譬如，穿越校园时，选择了这条路而不是另一条路。但这并不意味着这个行为有何错误。（这些事情简直太微不足道了。）哪怕我“任意地”做一件举足轻重的事情，譬如改换工作，如果我不承担家庭责任的话，也可以只是我的权利。但是如果这个行动要求提供正当理由，那么，任意地、无正当理由地去行动便要遭到合情合理的谴责。作出一个在很大程度上没有正当理由的、仅仅武断而任意的（又是一个道德描述词）对他人的生命（很可能还有他自己的生命）产生严重有害影响的决定，是不合理性的典型例子，不仅是不合理性，而且是邪恶的典型例子。

在第七章，我们先讨论了边沁的论点：“撇开偏见不论”，针戏游戏（一种古代的儿童游戏）和“艺术，音乐和诗”有同等价值。在边沁看来，诗确实比针戏更好的唯一理由，说穿了，只是诗比针戏给人们以更大的满足（或给更多的人以满足，或这两者兼而有之）这一浅显的事实。从根本上说，这一观点有两点错误。一是满足（自我利益）本身不可能是一个人的唯一目标。假若我除了“我的福利”之外，别无其他目标，那么我的福利就会是一个无意义的概念，这个观点可追溯到巴特勒主教。更重要的是，某些满足比其他满足

更好，“更高尚”，且言之有据，诗和音乐给人以慰藉，它们开拓了我们的感受，它们给许多人（包括许多有史以来最有天才的大师）提供了一个自我表达的重要方式。把这些区分高级满足和低级满足的理由称为“偏见”，这种做法事实上和两要素论以及价值性质是“本体论上可疑的”这两种观点密切相关。边沁运用的模式正是“中性事实”和任意“偏见”的模式。的确，称对诗的爱好为一种偏见，这是边沁用来表达如下观点的独特方式，这种观点以为，诗比针戏给人以更大的满足这一事实是在比较两者的时候作出的唯一不是“任意的”考虑，而对某种满足而不是别种满足的任何爱好据说都是任意的。但这根本不对，只要考虑“任意的”爱好这一概念在我们的概念体系中的实际地位，上述观点便完全是错误的。而且，一旦“任意的”脱离了它所属的概念体系，这种观点还是毫无意义的。（同样地，说宁要诗而不要针戏是一种偏见的陈述，严格讲也是不能成立的。）现在有人正主张，承认有满足这种东西在本体论是合法的，而承认有扩充了的感受能力、指谓和隐喻的多种手法、表达方式和自我实现方式等东西，则是本体论上不合法的。价值不属于世界内容的观点，和“价值判断”是对“偏见”的表述的观点，是同一事物的两个方面。

我们已经考察了能否合理地支持“价值判断”的问题。我们已看到，各种否定的回答都建立在各种可疑的哲学假设之上。这些假设是：合理性本身只是对预测有用的；或者，仅仅对达到“同意”有利；或者，达到真理只有一种方法（有时，在这种观点看来，“真理”的唯一标准据说就是预测和同意）；或者，价值判断有“两个意义要素”；或者，价值性质在“本体论上是可疑的”。我所辩护的立场是：任何概念系统的选择都预设了价值，而且，为描述日常的人际关系和社会事实（更不要说考虑到某人自身的生活理想）而进行的概念系统的选择，所包含的东西之一，便是一个人的道德价值。一个人不能选择只是事实的“摹本”的系统，因为没有一个概念系统是世

界的单纯“摹本”。真理概念本身就其内容而言就依赖我们的合理的可接受性标准，而这些合理的可接受性标准则依赖于并预设了我们的价值。更简单地说，真理论以合理性理论为前提，合理性理论又以我们关于善的理论为前提。

但是，“善的理论”不单单是带有纲领性的理论，它本身还得依赖关于人性、社会和宇宙的种种假定（包括神学和形而上学的假定）。随着我们知识的增长，世界观的变化，我们事实上不得不一再修改我们的善的理论。

事实很清楚，在我们辩护的概念中，根本没有“基础”这种东西存在。就这点而论，人们会担忧：如此一来，我们不是朝着抹煞“有正当理由”和“在我们看来有正当理由”（相对主义）之间的区别，甚至和“在我个人看来有正当理由”（唯我论之一种）之间的区别这种观点靠拢了吗？

只要我们极力站在合理性所属的概念体系之外，同时又自诩提供了关于合理性的更“合理”概念，那么唯我论就确是我们所将要达到的归宿！（许多思想家都陷入了尼采式的错误，宣称他们提出了优于全部传统所曾有过的道德观念。但每次都只能产生一个怪物，因为他们所做的，不外乎任意地把某些价值从其历史背景中分离出来，同时对其他道德视而不见。）我们只能希望在我们自身的传统（这个传统，在合理性方面，使人联想到古希腊广场辩论和牛顿的理论；在道德方面，使人联想到《圣经》，哲学家和民主革命）之内产生一个更为合理的合理性观念，或更好的道德观念。但这决不是说，我们现在所有的观念都是完全合理的，完美无缺的。我们没有落入个体唯我论的地狱，而是被引向一个真正的人类对话之中。这是一个将集体性和个人责任性融为一体的对话。

这个对话有一个理想终点吗？有没有一个真的合理性观念，一种真的道德，即使我们所曾有的都是我们对于这些东西的观念？在这里，哲学家们就像其他任何人一样分道扬镳了。罗蒂在他的美国

哲学学会主席的就职致辞[1]中竭力主张，唯一存在的只有对话，不可能设立、也不应该要求任何理想目标。但是"唯一存在的只有对话"这个断言怎样才能与我们在第五章中讨论的自相矛盾的相对主义相区分呢？我们把我们的不同观念看作是对合理性的不同观念来谈论这一事实本身，便设定了一个极限观念，一个理想真理的极限观念。

① "实用主义、相对主义和非理性主义"，见《美国哲学学会致辞及论文汇编》，1980年8月。也可见罗蒂的《哲学和自然之镜》，普林斯顿大学出版社，1979年版。

附　　录

以下便是第二章提到的那个定理。

定理　令 L 为一个具有 $F_1, F_2, \cdots, F_k$ 这些谓词（不一定是一元的）的语言。令 I 为 L 的诠释，即对 L 的每一个谓词指定一个内涵。这样，如果 I 是有意义的，也就是说，如果至少一个谓词在至少一个可能世界中具有一个既非空类亦非全类的外延，那么，存在着另一个诠释 J，它与 I 不一致，但像 I 一样使同一些句子在每一个可能世界中为真。

证明：令 $W_1, W_2, \cdots$ 为一切具有某种良序的可能世界，并且令 U_i 为存在于世界 W_i 中的可能个体的集。令 R_{ij} 为由 I 规定的可能世界 W_j 中谓词 F_i 的外延集（如果 F_{ij} 是非一元的，那么 R_{ij} 将是一个 Ni 重的集，其 N_i 为 F_i 的主目位置的数目）。结构〈U_j；R_{ij}（$i=1, 2, \cdots, k$）〉是世界 W_j 中相关于 I 的 L 的“预期模型”（亦即，U_j 是 L 在世界 W_j 中的论域），而（对于 $i=1, 2, \cdots, k$ 来说）R_{ij} 是 W_j 中谓词的 F_i 的外延。

如果至少有一个谓词（比如 F_u）具有一个既非空类亦非全部 U_j 的外延 R_{uj}，那么，选取 U_j 的一个置换式 P_j，使得 $P_j(R_{uj}) \neq R_{uj}$。否则的话，令 P_j 为一恒等式。既然 P_j 是一置换式，结构〈U_j；$P_j(R_{ij})$（$i=1, 2, \cdots, k$）〉与〈U_j；R_{ij}（$i=1, 2, \cdots, k$）〉同构，因而成为 L 的同样一些句子的模型（亦即，成为 L 中那些根据 I 在 W_j 是真的句子的模型）。

令J为L的这样的诠释，它赋予谓词$F_i(i=1,2,\cdots,k)$以下内涵：其值在任何可能世界W_i中都是$P_j(R_{ij})$的函项$f_i(W)$，换言之，F_i在各个W_j的外延根据诠释J被定义为$P_j(R_{ij})$。既然$\langle U_j;P_j(R_{ij})(i=1,2,\cdots,k)\rangle$和$\langle U_j;R_{ij}(i=1,2,\cdots,k)\rangle$是同一个语句集的模型（因为它们具有同构性），所以同一些语句在各个可能世界中根据J是真的，根据I也是真的，而J在任何其中至少一个谓词有一实义外延的世界中，都与I不同。证毕。

注一：如果在一特定世界W_j中有两个不交叉的集合，它们是由I规定的、L的谓词在W_j中的外延——比方说，猫的集合和狗的集合——那么，如果存在着多于猫的狗（或者，存在着至少与狗同样多的猫）的话，我们就可以选取任何同猫的全集同样大小的狗的集合（或者，选取任何同狗的全集同样大小的猫的集合），并且选择一个P_j，它将所选取的狗的集合绘制在猫的全集上（或者，把所选取的猫的集合绘制在狗的全集上），反过来又将猫的全集绘制在所选取的狗的集合上（或者，将狗的全集绘制在所选取的猫的集合上）；这将确保，根据J，第一个谓词（其外延根据I是猫的全集）是根据J在W_i中狗的一个集合，或者，将确保第二个谓词（其外延根据I是狗的全集）是根据J在W_j中猫的一个集合。

注二：有些对象，人们不希望加以置换，因为人们在某种意义上把这些对象的谓词看作是“绝对的”。如果存在着这样的对象——比方说“感觉”——的话，人们干脆规定这些置换式P_j在这些对象方面是恒等式。这将造成这样的结果：将L中的任何谓词限于那些在任何世界中根据I和J都相同的特许对象。

注三：既然根据I和J，一些句子具有逻辑上等价的真值条

件，由此可知，根据这个标准的"可能世界语义学"，反事实条件句也成立。

人名中西文对照

阿尔图塞	L. Althusser
阿佩尔	K. O. Apel
艾克尔斯	J. Eccles
爱因斯坦	Einstein
柏拉图	Plato
贝克莱	Berkeley
贝克	J. Baker
边沁	Bentham
伯雷特	C. Burret
波普尔	K. Popper
博伊德	R. Boyd
波义耳	R. Boyle
伯克斯	A. Burks
布洛克	N. Block
达尔文	Darwin
达·芬奇	Leonardo Da Vinci
达米特	M. Dummett
戴维森	D. Davidson
丹尼特	D. Dennett
费尼提	D. Finetti
狄德罗	Diderot
笛卡儿	Descartes
杜威	J. Dewey

法伊阿本德	P. Feyerabend
费尔德	H. Field
弗雷格	G. Frege
弗洛伊德	Freud
福科	M. Foucault
福特	P. Foot
格莱斯	P. Grice
格里芬	D. Griffin
格列莫尔	C. Glymour
古德曼	N. Goodman
哈贝马斯	J. Habermas
哈勒	R. Harre
哈曼	G. Harman
黑格尔	Hegel
胡塞尔	E. Husserl
凯恩斯	J. M. Keynes
康德	Kant
刘易斯	C. S. Lewis
苛勒	P. Kolers
柯勒	W. Kohler
克里普克	S. Kripke
孔德	Comte
库恩	T. Kuhn
蒯因	W. Quine
莱布尼兹	Leibniz
列宁	Lenin
罗蒂	R. Rorty
罗素	B. Russell
洛克	Locke
马赫	Mach

马克思	Marx
马拉蒙特	D. Malament
麦基	J. Mackie
摩尔	G. E. Moore
穆勒	John Stuart Mill
默多克	I. Murdoch
内格尔	T. Nagel
尼采	Nietzsche
牛顿	Newton
纽拉特	O. Neurath
纽曼	John Newman
培根	Francis Bacon
皮尔士	C. S. Pierce
普特南	R. A. Putnam
卡尔纳普	R. Carnap
卡维尔	S. Cavell
乔姆斯基	Chomsky
加芬克尔	A. Garfinkel
丘吉尔	Churchill
斯宾诺莎	Spinoza
史蒂文森	C. Stevenson
塔斯基	A. Tarski
图灵	A. Turing
威廉斯	B. A. O. Williams
韦伯	M. Weber
维特根斯坦	L. Wittgenstein
休谟	Hume
亚里士多德	Aristotle
韦金斯	D. Wiggins
泽马赫	E. Zamach

恩皮里柯　Sextus Empiricus
莫诺　J. Monod
斯马特　J. J. C. Smart
萨维奇　L. J. Savage
摩西　Moses
蒙塔格　R. Montague
诺齐克　R. Nozick